(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 November 2002 (14.11.2002)

PCT

(10) International Publication Number
WO 02/090580 A1

(51) International Patent Classification[7]: C12Q 1/68

(21) International Application Number: PCT/AU02/00558

(22) International Filing Date: 3 May 2002 (03.05.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/288,441 3 May 2001 (03.05.2001) US

(71) Applicants *(for all designated States except US)*: **NATIONAL CANCER CENTRE OF SINGAPORE PTE LTD** [SG/SG]; 11 Hospital Drive, 169610 Singapore (SG). **SINGAPORE GENERAL HOSPITAL PTE LTD** [SG/SG]; 11 Third Hospital, SNEC Building, 7th Floor, 168751 Singapore (SG).

(71) Applicant *(for AU only)*: **ARGAET, Victor, Peter** [AU/AU]; 29 Beaconsfield Terrace, Gordon Park, Queensland 4031 (AU).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **HUYNH, The Hung** [CA/SG]; 1 Pearlbank #26-03, 169016 Singapore (SG). **CHOW, Pierce, Kah, Hoe** [MY/SG]; 10A Siak Kew Avenue, 348081 Singapore (SG). **SOO, Khee, Chee** [MY/SG]; 4 Tan Boon Chong Avenue, 1027 Singapore (SG).

(74) Agents: **ARGAET, Victor, Peter** et al.; Davies Collison Cave, Level 3, 303 Coronation Drive, Milton, Queensland 4064 (AU).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 02/090580 A1

(54) Title: IDENTIFYING LIVER CANCER BY DETECTING ABERRANT EXPRESSION OF INSULIN LIKE GROWTH FACTOR BINDING PROTEIN

(57) Abstract: The invention is the linkage of the aberrant expression of insulin like growth factor binding proteins (IGFBPs) to the development and/or progression of liver cancer. It includes a method for detecting the presence or diagnosing the risk of a liver cancer, either before or after the onset of clinical symptoms, by detecting an aberrant level and/or functional activity of an expression product of a gene encoding an IGFBP, which correlates with the presence or risk of liver cancer. Also disclosed are methods of screening for agents that modulate the level and/or functional activity of an expression product of an IGFBP gene or gene belonging to the same biosynthetic or regulatory pathway as an IGFBP gene for use in methods for modulating liver cell proliferation of tumourgeneis, or for treating and/or preventing a liver cancer or related condition.

IDENTIFYING LIVER CANCER BY DETECTING ABERRANT EXPRESSION OF INSULIN LIKE GROWTH FACTOR BINDING PROTEIN

FIELD OF THE INVENTION

THIS INVENTION relates generally to polynucleotides and polypeptides linked to cancer. More particularly, the present invention relates to the aberrant expression of insulin-like growth factor binding proteins (IGFBPs), which is linked to the development and/or progression of cancer, especially a cancer of the liver such as hepatocellular carcinoma (HCC). Even more particularly, the present invention relates to a method for detecting the presence or diagnosing the risk of a liver cancer, either before or after the onset of clinical symptoms, by detecting an aberrant level and/or functional activity of an expression product of a gene encoding an insulin-like growth factor binding protein (IGFBP), which correlates with the presence or risk of liver cancer. Assessing the level and/or functional activity of the aberrant expression product is useful as a prognostic indicator of disease outcome. The invention also encompasses methods of screening for agents that modulate the level and/or functional activity of an expression product of an *IGFBP* gene or a gene belonging to the same biosynthetic or regulatory pathway as an *IGFBP* gene for use in methods for modulating liver cell proliferation or tumorigenesis, of for treating and/or preventing a liver cancer or related condition.

Bibliographic details of the publications numerically referred to in this specification are collected at the end of the description.

BACKGROUND OF THE INVENTION

Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide but treatment outcomes for HCC have remained generally poor. The majority of patients with HCC have inoperable disease with very poor prognosis (1). Survival in patients with curative resection carried out at dedicated centres is between 35 - 50% at 5 years and much lower elsewhere (2, 3). Long-term survival is uncommon because of the frequent presence of recurrence, metastasis or the development of new primaries (4, 5). There is also currently no accepted adjuvant or palliative treatment modalities that have been conclusively shown to prolong survival in HCC (6).

Recent research has shown that insulin-like growth factors I and II (IGF-I and IGF-II) are potent mitogens for human hepatoma cells (7), and that both IGF-I mRNA (8) and IGF-I receptors (IGF-IR) (7, 8) are detectable in human hepatoma cell lines. Both IGF-

I and IGF-II bind with high affinity to specific IGF-binding proteins (IGFBPs), which modulate their bioactivity. At least six IGFBPs have been described (reviewed in (9, 10)). Expression of genes encoding the various IGFBPs has been observed in many tissues, and is subject to intricate physiological regulation (reviewed in (9, 10)). IGFBP-3 is the most abundant IGF binding protein in the circulation, where it forms a 150 kDa complex with an acid-labile subunit and IGF-I or IGF-II (10). The IGFBP-3 gene is expressed in many tissues and IGFBP-3 has affinities for IGFs that are either equal to or stronger than those of the IGF receptors and therefore inhibits the IGFs by sequestration in the extracellular compartment (reviewed in (9, 10)). Recent evidence also demonstrates that IGFBP-3 has growth inhibitory activity that is independent of its IGF binding properties (reviewed in (10)).

Existing evidence suggests that early diagnosis of HCC may allow modulation of its activities and the potential for therapeutical applications in the control of HCC. Conventional methods for diagnosis of HCC are based on measuring serum levels of alpha-fetoprotein (AFP). Normally, AFP is a major serum protein synthesised by foetal liver cells, yolk sac cells, and in trace amounts by the foetal gastrointestinal tract. Reappearance of AFP in adult serum often signals pathologic conditions, particularly the presence of HCC and germ cell tumours containing yolk sac cell elements. Although existing assays may be used successfully for monitoring treatment of AFP-producing tumours and as an independent prognostic tool, the finding of elevated serum AFP levels in some patients with nonmalignant liver diseases, particularly in acute and chronic viral hepatitis and cirrhosis, has limited the value of such assays as an independent specific test to establish the diagnosis of cancer.

An acute need, therefore, still exists in the art for an assay which is selective for patients with HCC, particularly given the need for early detection of HCC in high risk populations.

SUMMARY OF THE INVENTION

The present invention is predicated in part on the determination that IGFBPs are differentially expressed in liver cancer (*e.g.*, HCC). Thus, IGFBPs and the transcripts from which they are translated have utility, for example, as markers of liver cancer, and as targets for therapeutic intervention in liver cancer. Accordingly, in one aspect of the present invention, there is provided a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient aberrant expression of a gene encoding an insulin-like growth factor binding protein (IGFBP). Preferably, the IGFBP is selected from IGFBP–1, IGFBP–2 or IGFBP–3. The liver cancer is preferably a liver cell carcinoma.

In another aspect, the invention contemplates a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient aberrant expression of at least one gene encoding an IGFBP selected from IGFBP–1, –2 and –3.

In yet another aspect, the invention encompasses a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient a level and/or functional activity of an expression product of a gene encoding an IGFBP–1, which is elevated relative to a normal reference level and/or functional activity of said expression product.

In still yet another aspect, the invention envisions a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient a level and/or functional activity of an expression product of a gene encoding an IGFBP–2, which is elevated relative to a normal reference level and/or functional activity of said expression product.

In a further aspect, the invention contemplates a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient a level and/or functional activity of an expression product of a gene encoding an IGFBP–3, which is reduced relative to a normal reference level and/or functional activity of said expression product.

In yet a further aspect, the invention resides in a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and/or

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and/or

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

In another aspect, the invention contemplates a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product.

In yet another aspect, the invention encompasses a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

In yet another aspect, the invention envisions a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

In still yet another aspect, the invention features a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

In another aspect, the invention provides a method for diagnosis in a patient of the progression of a liver cancer, comprising detecting aberrant expression of at least one gene encoding an IGFBP in a biological sample obtained from said patient.

In yet another aspect, the invention contemplates a method for prognostic assessment of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient aberrant expression of at least one gene encoding an IGFBP.

In yet another aspect, the invention extends to the use of an agent in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of a gene selected from an *IGFBP* gene or a gene relating to the same regulatory or biosynthetic pathway as the *IGFBP* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a change in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a change in the level of expression of said marker gene.

In another aspect, the invention contemplates the use of an agent that reduces the level and/or functional activity of an expression product of a gene encoding an IGFBP—1 in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–1* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

In yet another aspect, the invention provides the use of an agent that reduces the level and/or functional activity of an expression product of a gene encoding an IGFBP—2 in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–2* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

In a further aspect, the invention extends to the use of an agent that increases the level and/or functional activity of an expression product of a gene encoding an IGFBP—3 in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–3* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting an increase in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or an increase in the level and/or functional activity of an expression product of said marker gene.

In one embodiment, the agent is selected from (1) an antisense oligonucleotide or ribozyme that binds to, or otherwise interacts specifically with, a polynucleotide selected from an *IGFBP–1* polynucleotide, an *IGFBP–2* polynucleotide or complement of these, or (2) an antigen-binding molecule that is immuno-interactive with IGFBP–1 and/or IGFBP–2. In another embodiment, the agent is selected from a first member comprising an IGFBP–3 polypeptide or a biologically active fragment thereof, or variant or derivative of these or a second member comprising a polynucleotide encoding said first member, which is operably connected to a regulatory polynucleotide.

In yet another aspect, the invention contemplates the use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises at least a portion of an *IGFBP–3* polynucleotide operably linked, in the sense direction, to a regulatory polynucleotide.

In another aspect, the invention extends to the use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises an isolated *IGFBP–1* polynucleotide operably linked, in the antisense direction, to a regulatory polynucleotide.

In yet another aspect, the invention provides the use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises an isolated *IGFBP–2* polynucleotide operably linked, in the antisense direction, to a regulatory polynucleotide.

In still another aspect, the invention contemplates a method for modulating liver cell proliferation or tumorigenesis, comprising contacting a liver cell with an agent as

broadly described above for a time and under conditions sufficient to modulate the level and/or functional activity of an *IGFBP* expression product.

According to another aspect, the invention provides a method for treatment and/or prophylaxis of a liver cancer or related condition, comprising administering to a patient in need of such treatment an effective amount of a compound selected from an agent as broadly described above, an agent-containing medicament as broadly described above, or a vector-containing medicament as broadly described above.

The invention also contemplates, in another aspect, the use of an agent-containing medicament as broadly described above, or a vector-containing composition as broadly described above in the preparation of a medicament for treating a liver cancer or related condition.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. IGFBP–3 expression in adjacent normal liver tissue and HCC tumours. Tissue lysates from normal adjacent liver tissue and HCC tumours were analysed by Western blotting. Blots were incubated with anti-human IGFBP–3 and α-tubulin antibodies. Representative samples are shown in panels (A), (B) and (C). All normal adjacent liver tissues had high levels of IGFBP–3 while IGFBP–3 protein was either under-expressed or absent from HCC tumours. Note that the low molecular weight IGFBP–3 fragments were present in serum samples (see panel (C)).

Figure 2. Immunostaining of normal liver (A) and HCC tumours (B) for IGFBP–3. Normal adjacent liver tissue and HCC tumours were stained with anti-human IGFBP–3 as described in Example 1. Adjacent normal liver tissue, showing intense expression of normal cells for IGFBP–3 while very low staining signal was observed for the morphologically disorganised HCC cells. Representative staining are shown. (original magnification X 400).

Figure 3. Immunostaining of normal liver and HCC tumours for IGF–I and IGF–II. Normal adjacent liver tissue (AC) and HCC tumours (B,D) were stained with anti-human IGF–I (A and B) or anti-human IGF–II (C and D) as described in Example 1. Representative samples are shown. Normal adjacent liver tissue shows intense staining for IGF–I while low staining signals are seen for HCC cells. HCC tumours show more intense staining for IGF–II than normal adjacent tissue. (original magnification X 400).

Figure 4. IGF–I receptor expression in adjacent normal liver tissue and HCC tumours. Tissue lysates from normal adjacent liver tissue and HCC tumours were analysed by Western blotting as described in Example 1. Blots were incubated with anti-IGF-IR (A) and anti-α-tubulin (B) antibodies. Note that in certain pairs of samples, normal adjacent liver tissue had higher levels of IGF–IR than HCC tumours.

Figure 5. Effects of IGFs, IGFBP–3, and IGF–II antibody on HepG2 cell proliferation. Cells were cultured as described in Example 1. Cells were incubated with serum free MEM, and indicated reagents of IGF–I for 48 h. Cell number was determined as described (14). Means of triplicate experiments were plotted. Bars with different letters are significantly different from one another at ($p<0.01$, Mann Whitney U test). SE at each point was <15%. *A*, Effect of IGF–I on HepG2 proliferation. IGF–I significantly increased HepG2 cell number. *B*, Effect of IGFBP–3 on HepG2 cell proliferation. Growth of HepG2

cells was significantly inhibited by exogenous IGFBP–3. *C*, Effect of hrIGFBP–3 on IGF–I-induced HepG2 cell proliferation. Cells were incubated with either serum free medium or 25 ng/mL IGF–I in the presence of various concentrations of hrIGFBP–3 for 48 h. IGF–I-induced HepG2 proliferation was significantly attenuated by hrIGFBP–3. *D*, Effect of IGF–II and anti-IGF–II antibody on HepG2 cell proliferation. Cells were incubated with serum free, 25 ng/mL hrIGF–II, rabbit preimmune serum (1:800 dilution), rabbit anti-human IGF–II antibody (2 μg/mL) and 25 ng/mL IGF–I in combination with rabbit anti-human IGF–II (2 μg/mL). IGF–II-induced proliferation rates were significant decreased in the presence of anti-IGF–II antibody. Note that IGF–II antibody also inhibits HepG2 cell growth.

Figure 6. Serum levels of IGFBP–1, –2, –3 and alpha fetoprotein in non-HCC and HCC patients. Serum samples from non-HCC and HCC patients were analysed by Western blotting as described in Example 2. Blots were incubated with anti-human IGFBP–1, anti-bovine IGFBP–2, anti-human IGFBP–3 and anti-human alpha fetoprotein antibodies. Representative samples are shown in panels 6A – 6P. All non-HCC serum samples had high levels of IGFBP–3 while IGFBP–1 and IGFBP–2 were low. IGFBP–3 was either low or absent from HCC. Upregulation of IGFBP–1 and IGFBP–2 was observed in most HCC serum samples. Note that AFP was absent in some HCC serum samples.

BRIEF DESCRIPTION OF THE SEQUENCES: SUMMARY TABLE

TABLE A

SEQUENCE ID	*DESCRIPTION*	*LENGTH*
SEQ ID NO: 1	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 1 (IGFBP–1), mRNA as set forth in GenBank Accession No. XM_004688	1514 nts
SEQ ID NO: 2	IGFBP–1 polypeptide encoded by SEQ ID NO: 1	259 aa
SEQ ID NO: 3	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 2 (IGFBP–2), mRNA as set forth in GenBank Accession No. NM_000597	1433 nts
SEQ ID NO: 4	IGFBP–2 polypeptide encoded by SEQ ID NO: 3	328 aa
SEQ ID NO: 5	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 3 (IGFBP–3), mRNA as set forth in GenBank Accession No. BC000013	2497 nts
SEQ ID NO: 6.	IGFBP–3 polypeptide encoded by SEQ ID NO: 5	291 aa
SEQ ID NO: 7	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 1 (IGFBP–1) gene as set forth in GenBank Accession No. M74587	6128 nts
SEQ ID NO: 8	IGFBP–1 polypeptide encoded by SEQ ID NO: 7	259 aa
SEQ ID NO: 9	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 2 (IGFBP–2) gene as set forth in GenBank Accession No. gi|18552832:211494-242517	31024 nts
SEQ ID NO: 10	IGFBP–2 polypeptide encoded by SEQ ID NO: 7	328 aa

SEQUENCE ID	*DESCRIPTION*	*LENGTH*
SEQ ID NO: 11	Nucleotide sequence corresponding to human insulin-like growth factor binding protein 3 (IGFBP–3) gene as set forth in GenBank Accession No. M35878	10884 nts
SEQ ID NO: 12	IGFBP–3 polypeptide encoded by SEQ ID NO: 11	291 aa

DETAILED DESCRIPTION OF THE INVENTION

1. Definitions

Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by those of ordinary skill in the art to which the invention belongs. Although any methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present invention, preferred methods and materials are described. For the purposes of the present invention, the following terms are defined below.

The articles *"a"* and *"an"* are used herein to refer to one or to more than one (i.e. to at least one) of the grammatical object of the article. By way of example, "an element" means one element or more than one element.

The term *"aberrant expression product"* refers to expressed polynucleotides or polypeptide, which result from a substitution, deletion and/or addition of one or more nucleotides or amino acid residues in a "normal" reference polynucleotide or polypeptide, and which correlate with the presence or risk of a liver cancer.

"Amplification product" refers to a nucleic acid product generated by nucleic acid amplification techniques.

By *"antigen-binding molecule"* is meant a molecule that has binding affinity for a target antigen. It will be understood that this term extends to immunoglobulins, immunoglobulin fragments and non-immunoglobulin derived protein frameworks that exhibit antigen-binding activity.

The term *"biological sample"* as used herein refers to a sample that may be extracted, untreated, treated, diluted or concentrated from a patient. The biological sample may comprise a fluid selected from whole blood, serum, plasma, saliva, urine, sweat, ascitic fluid, peritoneal fluid, synovial fluid, amniotic fluid, cerebrospinal fluid, tissue biopsy (*e.g.*, liver biopsy), and the like. The biological sample preferably includes serum, whole blood, plasma and lymph or other circulatory fluid, saliva, mucus secretion and respiratory fluid. More preferably, the biological sample comprises a circulatory fluid such as serum or whole blood or a fractionated portion thereof. Most preferably, the biological sample comprises serum or a fractionated portion thereof.

By *"biologically active fragment"* is meant a fragment of a full-length parent polypeptide which fragment retains the activity of the parent polypeptide. A biologically active fragment will therefore have, for example, the activity of an IGFBP. As used herein, the term *"biologically active fragment"* includes deletion variants and small peptides, for example of at least 10, preferably at least 20 and more preferably at least 30 contiguous amino acids, which comprise the above activities. Peptides of this type may be obtained through the application of standard recombinant nucleic acid techniques or synthesised using conventional liquid or solid phase synthesis techniques. For example, reference may be made to solution synthesis or solid phase synthesis as described, for example, in Chapter 9 entitled *"Peptide Synthesis"* by Atherton and Shephard which is included in a publication entitled *"Synthetic Vaccines"* edited by Nicholson and published by Blackwell Scientific Publications. Alternatively, peptides can be produced by digestion of a polypeptide of the invention with proteinases such as endoLys-C, endoArg-C, endoGlu-C and staphylococcus V8-protease. The digested fragments can be purified by, for example, high performance liquid chromatographic (HPLC) techniques.

Throughout this specification, unless the context requires otherwise, the words *"comprise"*, *"comprises"* and *"comprising"* will be understood to imply the inclusion of a stated step or element or group of steps or elements but not the exclusion of any other step or element or group of steps or elements.

By "*effective amount*", in the context of treating or preventing a condition is meant the administration of that amount of active to an individual in need of such treatment or prophylaxis, either in a single dose or as part of a series, that is effective for the health and physical condition of the individual to be treated, the taxonomic group of individual to be treated, the formulation of the composition, the assessment of the medical situation, and other relevant factors. It is expected that the amount will fall in a relatively broad range that can be determined through routine trials.

As used herein, the term *"function"* refers to a biological, enzymatic, or therapeutic function.

Reference herein to *"immuno-interactive"* includes reference to any interaction, reaction, or other form of association between molecules and in particular where one of the molecules is, or mimics, a component of the immune system.

By *"isolated"* is meant material that is substantially or essentially free from components that normally accompany it in its native state.

By *"marker gene"* is meant a gene that imparts a distinct phenotype to cells expressing the marker gene and thus allows such transformed cells to be distinguished from cells that do not have the marker. A selectable marker gene confers a trait for which one can 'select' based on resistance to a selective agent (*e.g.*, antibiotic, radiation, heat, or other treatment damaging to untransformed cells). A screenable marker gene (or marker gene) confers a trait that one can identify through observation or testing, *i.e.*, by 'screening' (*e.g.* β-glucuronidase, luciferase, or other enzyme activity not present in untransformed cells).

By *"obtained from"* is meant that a sample such as, for example, a polynucleotide extract or polypeptide extract is isolated from, or derived from, a particular source of the host. For example, the extract can be obtained from a tissue or a biological fluid isolated directly from the host.

The term "*oligonucleotide*" as used herein refers to a polymer composed of a multiplicity of nucleotide residues (deoxyribonucleotides or ribonucleotides, or related structural variants or synthetic analogues thereof) linked via phosphodiester bonds (or related structural variants or synthetic analogues thereof). Thus, while the term "oligonucleotide" typically refers to a nucleotide polymer in which the nucleotide residues and linkages between them are naturally occurring, it will be understood that the term also includes within its scope various analogues including, but not restricted to, peptide nucleic acids (PNAs), phosphoramidates, phosphorothioates, methyl phosphonates, 2-O-methyl ribonucleic acids, and the like. The exact size of the molecule can vary depending on the particular application. An oligonucleotide is typically rather short in length, generally from about 10 to 30 nucleotide residues, but the term can refer to molecules of any length, although the term "polynucleotide" or "nucleic acid" is typically used for large oligonucleotides.

The term *"operably connected"* or *"operably linked"* as used herein means placing a structural gene under the regulatory control of a promoter, which then controls the transcription and optionally translation of the gene. In the construction of heterologous promoter/structural gene combinations, it is generally preferred to position the genetic sequence or promoter at a distance from the gene transcription start site that is approximately the same as the distance between that genetic sequence or promoter and the

gene it controls in its natural setting; *i.e.* the gene from which the genetic sequence or promoter is derived. As is known in the art, some variation in this distance can be accommodated without loss of function. Similarly, the preferred positioning of a regulatory sequence element with respect to a heterologous gene to be placed under its control is defined by the positioning of the element in its natural setting; *i.e.* the genes from which it is derived.

The term "*patient*" refers to patients of human or other mammal and includes any individual it is desired to examine or treat using the methods of the invention. However, it will be understood that "*patient*" does not imply that symptoms are present. Suitable mammals that fall within the scope of the invention include, but are not restricted to, primates, livestock animals (*eg*. sheep, cows, horses, donkeys, pigs), laboratory test animals (*eg*. rabbits, mice, rats, guinea pigs, hamsters), companion animals (*eg*. cats, dogs) and captive wild animals (*eg*. foxes, deer, dingoes).

By *"pharmaceutically acceptable carrier"* is meant a solid or liquid filler, diluent or encapsulating substance that can be safely used in topical or systemic administration to a animal, preferably a mammal including humans.

The term "*polynucleotide*" or "*nucleic acid*" as used herein designates mRNA, RNA, cRNA, cDNA or DNA. The term typically refers to oligonucleotides greater than 30 nucleotide residues in length. Polynucleotide sequences are understood to encompass complementary strands as well as alternative backbones described herein.

The terms *"polynucleotide variant"* and *"variant"* refer to polynucleotides displaying substantial sequence identity with a reference polynucleotide sequence or polynucleotides that hybridise with a reference sequence under stringent conditions that are defined hereinafter. These terms also encompasses polynucleotides in which one or more nucleotides have been added or deleted, or replaced with different nucleotides. In this regard, it is well understood in the art that certain alterations inclusive of mutations, additions, deletions and substitutions can be made to a reference polynucleotide whereby the altered polynucleotide retains the biological function or activity of the reference polynucleotide. The terms *"polynucleotide variant"* and *"variant"* also include naturally occurring allelic variants.

"*Polypeptide*", "*peptide*" and "*protein*" are used interchangeably herein to refer to a polymer of amino acid residues and to variants and synthetic analogues of the same. Thus, these terms apply to amino acid polymers in which one or more amino acid residues

is a synthetic non-naturally occurring amino acid, such as a chemical analogue of a corresponding naturally occurring amino acid, as well as to naturally-occurring amino acid polymers.

The term *"polypeptide variant"* refers to polypeptides in which one or more amino acids have been replaced by different amino acids. It is well understood in the art that some amino acids may be changed to others with broadly similar properties without changing the nature of the activity of the polypeptide (conservative substitutions).

By "*primer*" is meant an oligonucleotide which, when paired with a strand of DNA, is capable of initiating the synthesis of a primer extension product in the presence of a suitable polymerising agent. The primer is preferably single-stranded for maximum efficiency in amplification but can alternatively be double-stranded. A primer must be sufficiently long to prime the synthesis of extension products in the presence of the polymerisation agent. The length of the primer depends on many factors, including application, temperature to be employed, template reaction conditions, other reagents, and source of primers. For example, depending on the complexity of the target sequence, the oligonucleotide primer typically contains 15 to 35 or more nucleotide residues, although it can contain fewer nucleotide residues. Primers can be large polynucleotides, such as from about 200 nucleotide residues to several kilobases or more. Primers can be selected to be "substantially complementary" to the sequence on the template to which it is designed to hybridise and serve as a site for the initiation of synthesis. By "substantially complementary", it is meant that the primer is sufficiently complementary to hybridise with a target polynucleotide. Preferably, the primer contains no mismatches with the template to which it is designed to hybridise but this is not essential. For example, non-complementary nucleotide residues can be attached to the 5' end of the primer, with the remainder of the primer sequence being complementary to the template. Alternatively, non-complementary nucleotide residues or a stretch of non-complementary nucleotide residues can be interspersed into a primer, provided that the primer sequence has sufficient complementarity with the sequence of the template to hybridise therewith and thereby form a template for synthesis of the extension product of the primer.

"Probe" refers to a molecule that binds to a specific sequence or sub-sequence or other moiety of another molecule. Unless otherwise indicated, the term "probe" typically refers to a polynucleotide probe that binds to another polynucleotide, often called the "target polynucleotide", through complementary base pairing. Probes can bind target

polynucleotides lacking complete sequence complementarity with the probe, depending on the stringency of the hybridisation conditions. Probes can be labelled directly or indirectly.

The term *"recombinant polynucleotide"* as used herein refers to a polynucleotide formed *in vitro* by the manipulation of a polynucleotide into a form not normally found in nature. For example, the recombinant polynucleotide can be in the form of an expression vector. Generally, such expression vectors include transcriptional and translational regulatory polynucleotide operably linked to the polynucleotide.

By *"recombinant polypeptide"* is meant a polypeptide made using recombinant techniques, *ie.* through the expression of a recombinant or synthetic polynucleotide.

By *"reporter molecule"* as used in the present specification is meant a molecule that, by its chemical nature, provides an analytically identifiable signal that allows the detection of a complex comprising an antigen-binding molecule and its target antigen. The term "reporter molecule" also extends to use of cell agglutination or inhibition of agglutination such as red blood cells on latex beads, and the like.

By *"vector"* is meant a polynucleotide molecule, preferably a DNA molecule derived, for example, from a plasmid, bacteriophage, yeast or virus, into which a polynucleotide can be inserted or cloned. A vector preferably contains one or more unique restriction sites and can be capable of autonomous replication in a defined host cell including a target cell or tissue or a progenitor cell or tissue thereof, or be integrable with the genome of the defined host such that the cloned sequence is reproducible. Accordingly, the vector can be an autonomously replicating vector, *ie.* a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *eg.* a linear or closed circular plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector can contain any means for assuring self-replication. Alternatively, the vector can be one which, when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. A vector system can comprise a single vector or plasmid, two or more vectors or plasmids, which together contain the total DNA to be introduced into the genome of the host cell, or a transposon. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. In the present case, the vector is preferably a viral or viral-derived vector, which is operably functional in animal and preferably mammalian cells. Such vector may be derived from a poxvirus, an adenovirus or yeast. The vector can also include a selection

marker such as an antibiotic resistance gene that can be used for selection of suitable transformants. Examples of such resistance genes are known to those of skill in the art and include the *nptII* gene that confers resistance to the antibiotics kanamycin and G418 (Geneticin®) and the *hph* gene which confers resistance to the antibiotic hygromycin B.

As used herein, underscoring or italicising the name of a gene shall indicate the gene, in contrast to its protein product, which is indicated in the absence of any underscoring or italicising. For example, "*IGFBP–1*" shall mean the *IGFBP–1* gene, whereas "IGFBP–1" shall indicate the protein product of the "*IGFBP–1*" gene.

2. Methods of detecting aberrant IGFBP expression

The present invention is predicated in part on the discovery that the sera of patients with liver cancer, particularly hepatocellular carcinoma, have aberrant levels of insulin-like growth factor binding proteins (IGFBPs) relative to sera of normal patients. Thus, the invention features a method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting aberrant expression of a gene encoding an IGFBP in a biological sample obtained from said patient. The aberrant expression preferably relates to a level and/or functional activity of an expression product of said gene, which correlates with the presence or risk of said liver cancer. A correlation with the presence or risk of a liver cancer is made for example when the level and/or functional activity of an *IGFBP* expression product in the biological sample differs by at least 10%, more preferably at least 50%, even more preferably at least 100%, even more preferably at least 200%, even more preferably at least 400%, even more preferably at least 600% and still even more preferably at least 1000% compared to the level and/or functional activity of said expression product in a biological sample obtained from normal patients or from patients who are not afflicted with that liver cancer. In a preferred embodiment, the level and/or functional activity of an *IGFBP–1* or *IGFBP–2* expression product in said biological sample is at least 10%, more preferably at least 50%, even more preferably at least 100%, even more preferably at least 200%, even more preferably at least 400%, even more preferably at least 600% and still even more preferably at least 1000% higher than the level and/or functional activity of that expression product in a biological sample obtained from normal patients or from patients who are not afflicted with said liver cancer. In another preferred embodiment, the level and/or functional activity of an *IGFBP–3* expression product in said biological sample is at least 10%, more preferably at least 50%, even more preferably at least 100%, even more preferably at least 200%, even more

preferably at least 400%, even more preferably at least 600% and still even more preferably at least 1000% lower than the level and/or functional activity of that expression product in a biological sample obtained from normal patients or from patients who are not afflicted with said liver cancer.

Accordingly, the presence or risk of liver cancer is diagnosed when (1) an expression product of *IGFBP–1* or *IGFBP–2* is expressed at a higher level, or (2) an expression product of *IGFBP–3* is expressed at a lower level, compared to the level at which it is expressed in normal patients or in patients without liver cancer. In a preferred embodiment of this type, the method comprises detecting a level and/or functional activity of an expression product of an *IGFBP*-1 gene or an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of that expression product. In another preferred embodiment of this type, the method comprises detecting a level and/or functional activity of an expression product of an *IGFBP*-3 gene, which is reduced relative to a normal reference level and/or functional activity of that expression product.

Thus, it may be desirable for example to qualitatively or quantitatively determine IGFBP protein levels and/or *IGFBP* transcription levels. Alternatively or additionally, it may be desirable to search for aberrant *IGFBP* structural genes and regulatory regions, which encode aberrant *IGFBP* expression products or which produce aberrant levels of *IGFBP* expression products. Alternatively or additionally, it may be desirable to qualitatively or quantitatively determine the level of an expression product (*e.g.*, transcript, protein) of a gene relating to the same regulatory or biosynthetic pathway as an *IGFBP* gene, which can modulate or otherwise influence IGFBP protein levels and/or *IGFBP* transcription levels. Likewise, it may also be desirable to search for an aberrant gene relating to the same regulatory or biosynthetic pathway as an *IGFBP* gene.

The biological sample can be any tissue (*e.g.*, a liver biopsy) or fluid. The biological sample is preferably a biological fluid including, but not restricted to, a circulatory fluid such as whole blood, serum, plasma and lymph as well as other circulatory fluid and saliva, mucus secretion and respiratory fluid. More preferably, the biological sample is serum or a fractionated portion thereof.

2.1 Genetic Diagnosis

One embodiment of the instant invention comprises a method for detecting an increase in the expression of an *IGFBP* gene by qualitatively or quantitatively determining

the transcripts of an *IGFBP* gene in a cell (*e.g.*, a liver cell such a Kupffer cell). Exemplary human nucleic acid sequences for *IGFBP-1*, *IGFBP-2* and *IGFBP-3* mRNA and there corresponding genes are set forth in the enclosed Sequence Listing *infra* and are summarised in TABLE A *supra*.

Another embodiment of the instant invention comprises a method for detecting reduction in the expression or function of an *IGFBP* gene (*e.g.*, *IGFBP-3*), or enhancement of expression or function of an *IGFBP* gene (*e.g.*, *IGFBP-1, IGFBP-2*), by examining the genes and transcripts of a cell. Nucleic acid used is isolated from cells contained in the biological sample, according to standard methodologies (Sambrook, *et al.*, "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, 1989; Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons Inc, 1994-1998). The nucleic acid may be genomic DNA or fractionated or whole cell RNA. Where RNA is used, it may be desired to convert the RNA to a complementary DNA. In one embodiment, the RNA is whole cell RNA; in another, it is poly-A RNA. In one embodiment, the nucleic acid is amplified by a nucleic acid amplification technique. Suitable nucleic acid amplification techniques are well known to the skilled artisan, and include the polymerase chain reaction (PCR) as for example described in Ausubel *et al.* (*supra*); strand displacement amplification (SDA) as for example described in U.S. Patent No 5,422,252; rolling circle replication (RCR) as for example described in Liu *et al.*, (1996) and International application WO 92/01813) and Lizardi *et al.*, (International Application WO 97/19193); nucleic acid sequence-based amplification (NASBA) as for example described by Sooknanan *et al.*, (1994, *Biotechniques* **17**:1077-1080); and Q-β replicase amplification as for example described by Tyagi *et al.*, (1996, *Proc. Natl. Acad. Sci. USA* **93**: 5395-5400).

Depending on the format, the specific nucleic acid of interest is identified in the sample directly using amplification or with a second, known nucleic acid following amplification. Next, the identified product is detected. In certain applications, the detection may be performed by visual means (*e.g.*, ethidium bromide staining of a gel). Alternatively, the detection may involve indirect identification of the product via chemiluminescence, radioactive scintigraphy of radiolabel or fluorescent label or even via a system using electrical or thermal impulse signals (Affymax Technology; Bellus, 1994, J *Macromol. Sci. Pure, Appl. Chem.*, **A31**(1): 1355-1376).

Following detection, one may compare the results seen in a given patient with a control reaction or a statistically significant reference group of normal patients. In this

way, it is possible to correlate the amount of an IGFBP detected with the progression or severity of the disease.

In addition to determining levels of *IGFBP* transcripts, it also may prove useful to examine various types of defects. These defect could include deletions, insertions, point mutations and duplications. Point mutations result in stop codons, frameshift mutations or amino acid substitutions. Somatic mutations are those occurring in non-germline tissues. Germ-line tissue can occur in any tissue and are inherited. Mutations in and outside the coding region also may affect the amount of IGFBP produced, both by altering the transcription of the gene or in destabilising or otherwise altering the processing of either the transcript (mRNA) or protein.

A variety of different assays are contemplated in this regard, including but not limited to, fluorescent *in situ* hybridisation (FISH), direct DNA sequencing, pulse field gel electrophoresis (PFGE) analysis, Southern or Northern blotting, single-stranded conformation analysis (SSCA), RNase protection assay, allele-specific oligonucleotide (ASO), dot blot analysis, denaturing gradient gel electrophoresis, RFLP and PCR-SSCP.

2.1.1 Primers and Probes

Primers may be provided in double-stranded or single-stranded form, although the single-stranded form is preferred. Probes, while perhaps capable of priming, are designed to bind to a target DNA or RNA and need not be used in an amplification process. In preferred embodiments, the probes or primers are labelled with radioactive species ^{32}P, ^{14}C, ^{35}S, ^{3}H, or other label), with a fluorophore (rhodamine, fluorescein) or a chemillumiscent label (luciferase).

2.1.2 Template Dependent Amplification Methods

A number of template dependent processes are available to amplify the target sequences present in a given template sample. An exemplary nucleic acid amplification technique is the polymerase chain reaction (referred to as PCR) which is described in detail in U.S. Pat. Nos. 4,683,195, 4,683,202 and 4,800,159, Ausubel *et al.* (*supra*), and in Innis *et al.*, ("PCR Protocols", Academic Press, Inc., San Diego Calif., 1990).

Briefly, in PCR, two primer sequences are prepared that are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates are added to a reaction mixture along with a DNA polymerase, *e.g.*, Taq polymerase. If the target sequence is present in a sample, the primers

will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction products and the process is repeated.

A reverse transcriptase PCR amplification procedure may be performed in order to quantify the amount of mRNA amplified. Methods of reverse transcribing RNA into cDNA are well known and described in Sambrook *et al.*, 1989. Alternative methods for reverse transcription utilise thermostable, RNA-dependent DNA polymerases. These methods are described in WO 90/07641. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction ("LCR"), disclosed in EPO No. 320 308. In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Pat. No. 4,883,750 describes a method similar to LCR for binding probe pairs to a target sequence.

Qβ Replicase, described in PCT Application No. PCT/US87/00880, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'α-thio-triphosphates in one strand of a restriction site may also be useful in the amplification of nucleic acids in the present invention, Walker *et al.*, (1992, Proc. Natl. Acad. Sci. U.S.A **89**: 392-396).

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, i.e., nick translation. A similar method, called Repair Chain Reaction (RCR), involves annealing several probes throughout a region targeted for

amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA. Target specific sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having 3' and 5' sequences of non-specific DNA and a middle sequence of specific RNA is hybridised to DNA that is present in a sample. Upon hybridisation, the reaction is treated with RNase H, and the products of the probe identified as distinctive products that are released after digestion. The original template is annealed to another cycling probe and the reaction is repeated.

Still another amplification method, described in GB Application No. 2 202 328, and in PCT Application No. PCT/US89/01025, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template- and enzyme-dependent synthesis. The primers may be modified by labelling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labelled probes are added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labelled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS), including nucleic acid sequence based amplification (NASBA) and 3SR (Kwoh *et al.*, 1989, *Proc. Natl. Acad. Sci. U.S.A.*, **86**: 1173; Gingeras *et al.*, PCT Application WO 88/10315). In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a clinical sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer which has target specific sequences. Following polymerisation, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target specific primer, followed by polymerisation. The double-stranded DNA molecules are then multiply transcribed by an RNA polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into single stranded DNA, which is then converted to double stranded DNA, and then transcribed once again with an RNA polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target specific sequences.

Davey *et al.*, EPO No. 329 822 disclose a nucleic acid amplification process involving cyclically synthesising single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from the resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in duplex with either DNA or RNA). The resultant ssDNA is a template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to the template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of E. coli DNA polymerase I), resulting in a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

Miller *et al.* in PCT Application WO 89/06700 disclose a nucleic acid sequence amplification scheme based on the hybridisation of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic, *i.e.*, new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" and "one-sided PCR" (Frohman, M. A., In: "PCR Protocols: A Guide to Methods and Applications", Academic Press, N.Y., 1990; Ohara *et al.*, 1989, *Proc. Natl Acad. Sci. U.S.A.*, **86**: 5673-567).

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide, may also be used in the amplification step of the present invention. Wu *et al.*, (1989, *Genomics* **4**: 560).

2.1.3 Southern/Northern Blotting

Blotting techniques are well known to those of skill in the art. Southern blotting involves the use of DNA as a target, whereas Northern blotting involves the use of RNA as

a target. Each provide different types of information, although cDNA blotting is analogous, in many aspects, to blotting or RNA species.

Briefly, a probe is used to target a DNA or RNA species that has been immobilized on a suitable matrix, often a filter of nitrocellulose. The different species should be spatially separated to facilitate analysis. This often is accomplished by gel electrophoresis of nucleic acid species followed by "blotting" on to the filter.

Subsequently, the blotted target is incubated with a probe (usually labelled) under conditions that promote denaturation and rehybridization. Because the probe is designed to base pair with the target, the probe will binding a portion of the target sequence under renaturing conditions. Unbound probe is then removed, and detection is accomplished as described above.

2.1.4 Detection Methods

Products may be visualised in order to confirm amplification of the target sequences. One typical visualisation method involves staining of a gel with ethidium bromide and visualisation under UV light. Alternatively, if the amplification products are integrally labelled with radio- or fluorometrically-labelled nucleotides, the amplification products can then be exposed to x-ray film or visualised under the appropriate stimulating spectra, following separation.

In one embodiment, visualisation is achieved indirectly. Following separation of amplification products, a labelled nucleic acid probe is brought into contact with the amplified target sequence. The probe preferably is conjugated to a chromophore but may be radiolabelled. In another embodiment, the probe is conjugated to a binding partner, such as an antibody or biotin, and the other member of the binding pair carries a detectable moiety or reporter molecule.

In one embodiment, detection is by a labelled probe. The techniques involved are well known to those of skill in the art and can be found in many standard texts on molecular protocols. See Sambrook *et al.*, 1989. For example, chromophore or radiolabel probes or primers identify the target during or following amplification.

One example of the foregoing is described in U.S. Pat. No. 5,279,721, which discloses an apparatus and method for the automated electrophoresis and transfer of nucleic acids. The apparatus permits electrophoresis and blotting without external

manipulation of the gel and is ideally suited to carrying out methods according to the present invention.

In addition, the amplification products described above may be subjected to sequence analysis to identify specific kinds of variations using standard sequence analysis techniques. Within certain methods, exhaustive analysis of genes is carried out by sequence analysis using primer sets designed for optimal sequencing (Pignon *et al.*, 1994, *Hum. Mutat.* **3**: 126-132). The present invention provides methods by which any or all of these types of analyses may be used. Using, for example, the sequences set forth herein, oligonucleotide primers may be designed to permit the amplification of sequences throughout *IGFBP-1*, *-2*, and *–3* genes that may then be analysed by direct sequencing.

2.1.5 Kit Components

All the essential materials and reagents required for detecting and sequencing *IGFBP* genes and variants thereof may be assembled together in a kit. The kits may also optionally include appropriate reagents for detection of labels, positive and negative controls, washing solutions, dilution buffers and the like. For example, a nucleic acid-based detection kit may include (i) a polynucleotide according to the invention (which may be used as a positive control), (ii) an oligonucleotide primer according to the invention. Also included may be enzymes suitable for amplifying nucleic acids including various polymerases (Reverse Transcriptase, Taq, Sequenase™ DNA ligase etc. depending on the nucleic acid amplification technique employed), deoxynucleotides and buffers to provide the necessary reaction mixture for amplification. Such kits also generally will comprise, in suitable means, distinct containers for each individual reagent and enzyme as well as for each primer or probe. The kit may also contain instructions for performing the method of the invention.

2.1.6 Chip Technologies

Also contemplated by the present invention are chip-based DNA technologies such as those described by Hacia *et al.* (1996, *Nature Genetics* **14**: 441-447) and Shoemaker *et al.* (1996, *Nature Genetics* **14**: 450-456). Briefly, these techniques involve quantitative methods for analysing large numbers of genes rapidly and accurately. By tagging genes with oligonucleotides or using fixed probe arrays, one can employ chip technology to segregate target molecules as high density arrays and screen these molecules

on the basis of hybridisation. See also Pease *et al.* (1994, *Proc. Natl. Acad. Sci. U.S.A.* **91**: 5022-5026); Fodor *et al.* (1991, *Science* **251**: 767-773).

2.2 Protein-based diagnostics

2.2.1 Antigen-binding molecules

Antigen-binding molecules that are immuno-interactive with a target molecule of the present invention can be used in measuring an increase or decrease in *IGFBP* expression. Thus, the present invention also contemplates antigen-binding molecules that bind specifically to IGFBP–1, IGFBP–2 or IGFBP–3 polypeptides or to proteins that regulate or otherwise influence the level and/or functional activity of one or more of the aforesaid IGFBP polypeptides. For example, the antigen-binding molecules may comprise whole polyclonal antibodies. Such antibodies may be prepared, for example, by injecting a target molecule of the invention into a production species, which may include mice or rabbits, to obtain polyclonal antisera. Methods of producing polyclonal antibodies are well known to those skilled in the art. Exemplary protocols which may be used are described for example in Coligan *et al.*, "Current Protocols In Immunology", (John Wiley & Sons, Inc, 1991), and Ausubel *et al.*, (1994-1998, *supra*), in particular Section III of Chapter 11.

In lieu of the polyclonal antisera obtained in the production species, monoclonal antibodies may be produced using the standard method as described, for example, by Köhler and Milstein (1975, *Nature* **256**, 495-497), or by more recent modifications thereof as described, for example, in Coligan *et al.*, (1991, *supra*) by immortalising spleen or other antibody-producing cells derived from a production species which has been inoculated with target molecule of the invention. A number of human and other mammalian antibodies to IGFBP-1, -2 and -3 are known, and are available either commercially or through techniques well known to the art and industry.

The invention also contemplates as antigen-binding molecules Fv, Fab, Fab' and $F(ab')_2$ immunoglobulin fragments. Alternatively, the antigen-binding molecule may comprise a synthetic stabilised Fv fragment. Exemplary fragments of this type include single chain Fv fragments (sFv, frequently termed scFv) in which a peptide linker is used to bridge the N terminus or C terminus of a V_H domain with the C terminus or N-terminus, respectively, of a V_L domain. ScFv lack all constant parts of whole antibodies and are not able to activate complement. Suitable peptide linkers for joining the V_H and V_L domains are those which allow the V_H and V_L domains to fold into a single polypeptide chain

having an antigen binding site with a three dimensional structure similar to that of the antigen binding site of a whole antibody from which the Fv fragment is derived. Linkers having the desired properties may be obtained by the method disclosed in U.S. Patent No 4,946,778. However, in some cases a linker is absent. ScFvs may be prepared, for example, in accordance with methods outlined in Kreber *et al* (Kreber *et al.* 1997, *J. Immunol. Methods*; **201**(1): 35-55). Alternatively, they may be prepared by methods described in U.S. Patent No 5,091,513, European Patent No 239,400 or the articles by Winter and Milstein (1991, *Nature* **349**:293) and Plünckthun *et al* (1996, In *Antibody engineering: A practical approach*. 203-252).

Alternatively, the synthetic stabilised Fv fragment comprises a disulphide stabilised Fv (dsFv) in which cysteine residues are introduced into the V_H and V_L domains such that in the fully folded Fv molecule the two residues will form a disulphide bond therebetween. Suitable methods of producing dsFv are described for example in (Glockscuther *et al. Biochem.* **29**: 1363-1367; Reiter *et al.* 1994, *J. Biol. Chem.* **269**: 18327-18331; Reiter *et al.* 1994, *Biochem. 33*: 5451-5459; Reiter *et al.* 1994. *Cancer Res.* **54**: 2714-2718; Webber *et al.* 1995, *Mol. Immunol.* **32**: 249-258).

Also contemplated as antigen-binding molecules are single variable region domains (termed dAbs) as for example disclosed in (Ward *et al.* 1989, *Nature* **341**: 544-546; Hamers-Casterman *et al.* 1993, *Nature*. **363**: 446-448; Davies & Riechmann, 1994, *FEBS Lett.* **339**: 285-290).

Alternatively, the antigen-binding molecule may comprise a "minibody". In this regard, minibodies are small versions of whole antibodies, which encode in a single chain the essential elements of a whole antibody. Suitably, the minibody is comprised of the V_H and V_L domains of a native antibody fused to the hinge region and CH3 domain of the immunoglobulin molecule as, for example, disclosed in U.S. Patent No 5,837,821.

In an alternate embodiment, the antigen binding molecule may comprise non-immunoglobulin derived, protein frameworks. For example, reference may be made to (Ku & Schultz, 1995, *Proc. Natl. Acad. Sci. USA*, **92**: 652-6556) which discloses a four-helix bundle protein cytochrome b562 having two loops randomised to create complementarity determining regions (CDRs), which have been selected for antigen binding.

The antigen-binding molecule may be multivalent (*ie.* having more than one antigen-binding site). Such multivalent molecules may be specific for one or more antigens. Multivalent molecules of this type may be prepared by dimerisation of two

antibody fragments through a cysteinyl-containing peptide as, for example disclosed by (Adams *et al.*, 1993, *Cancer Res.* **53**: 4026-4034; Cumber *et al.*, 1992, *J. Immunol.* **149**: 120-126). Alternatively, dimerisation may be facilitated by fusion of the antibody fragments to amphiphilic helices that naturally dimerise (Pack P. Plünckthun, 1992, *Biochem.* **31**: 1579-1584), or by use of domains (such as the leucine zippers jun and fos) that preferentially heterodimerise (Kostelny *et al.*, 1992, *J. Immunol.* **148**: 1547-1553). In an alternate embodiment, the multivalent molecule may comprise a multivalent single chain antibody (multi-scFv) comprising at least two scFvs linked together by a peptide linker. In this regard, non-covalently or covalently linked scFv dimers termed "diabodies" may be used. Multi-scFvs may be bispecific or greater depending on the number of scFvs employed having different antigen binding specificities. Multi-scFvs may be prepared for example by methods disclosed in U.S. Patent No. 5,892,020.

The antigen-binding molecules of the invention may be used for affinity chromatography in isolating a natural or recombinant polypeptide or biologically active fragment of the invention. For example reference may be made to immunoaffinity chromatographic procedures described in Chapter 9.5 of Coligan *et al.*, (1995-1997, *supra*).

2.2.2 Immunodiagnostic assays

The above antigen-binding molecules have utility in measuring directly or indirectly modulation of *IGFBP* expression in healthy and diseased states, through techniques such as ELISAs and Western blotting. Illustrative assay strategies which can be used to detect a target polypeptide of the invention include, but are not limited to, immunoassays involving the binding of an antigen-binding molecule to the target polypeptide (*e.g.*, an IGFBP polypeptide) in the sample, and the detection of a complex comprising the antigen-binding molecule and the target polypeptide. Preferred immunoassays are those that can measure the level and/or functional activity of a target molecule of the invention. Typically, an antigen-binding molecule that is immuno-interactive with a target polypeptide of the invention is contacted with a biological sample suspected of containing said target polypeptide. The concentration of a complex comprising the antigen-binding molecule and the target polypeptide is measure in and the measured complex concentration is then related to the concentration of target polypeptide in the sample. Consistent with the present invention, the presence of an aberrant

concentration of the target polypeptide is indicative of the presence or risk of a liver cancer.

Any suitable technique for determining formation of an antigen-binding molecule-target antigen complex may be used. For example, an antigen-binding molecule according to the invention, having a reporter molecule associated therewith may be utilised in immunoassays. Such immunoassays include, but are not limited to, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs) and immunochromatographic techniques (ICTs), Western blotting which are well known those of skill in the art. For example, reference may be made to Coligan *et al.* (1994, *supra*) which discloses a variety of immunoassays that may be used in accordance with the present invention. Immunoassays may include competitive assays as understood in the art or as for example described *infra*. It will be understood that the present invention encompasses qualitative and quantitative immunoassays.

Suitable immunoassay techniques are described for example in US Patent Nos. 4,016,043, 4, 424,279 and 4,018,653. These include both single-site and two-site assays of the non-competitive types, as well as the traditional competitive binding assays. These assays also include direct binding of a labelled antigen-binding molecule to a target antigen.

Two site assays are particularly favoured for use in the present invention. A number of variations of these assays exist, all of which are intended to be encompassed by the present invention. Briefly, in a typical forward assay, an unlabelled antigen-binding molecule such as an unlabelled antibody is immobilised on a solid substrate and the sample to be tested brought into contact with the bound molecule. After a suitable period of incubation, for a period of time sufficient to allow formation of an antibody-antigen complex, another antigen-binding molecule, suitably a second antibody specific to the antigen, labelled with a reporter molecule capable of producing a detectable signal is then added and incubated, allowing time sufficient for the formation of another complex of antibody-antigen-labelled antibody. Any unreacted material is washed away and the presence of the antigen is determined by observation of a signal produced by the reporter molecule. The results may be either qualitative, by simple observation of the visible signal, or may be quantitated by comparing with a control sample containing known amounts of antigen. Variations on the forward assay include a simultaneous assay, in which both sample and labelled antibody are added simultaneously to the bound antibody. These

techniques are well known to those skilled in the art, including minor variations as will be readily apparent. In accordance with the present invention, the sample is one that might contain an antigen including a tissue or fluid as described above.

In the typical forward assay, a first antibody having specificity for the antigen or antigenic parts thereof is either covalently or passively bound to a solid surface. The solid surface is typically glass or a polymer, the most commonly used polymers being cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene. The solid supports may be in the form of tubes, beads, discs of microplates, or any other surface suitable for conducting an immunoassay. The binding processes are well known in the art and generally consist of cross-linking covalently binding or physically adsorbing, the polymer-antibody complex is washed in preparation for the test sample. An aliquot of the sample to be tested is then added to the solid phase complex and incubated for a period of time sufficient and under suitable conditions to allow binding of any antigen present to the antibody. Following the incubation period, the antigen-antibody complex is washed and dried and incubated with a second antibody specific for a portion of the antigen. The second antibody has generally a reporter molecule associated therewith that is used to indicate the binding of the second antibody to the antigen. The amount of labelled antibody that binds, as determined by the associated reporter molecule, is proportional to the amount of antigen bound to the immobilised first antibody.

An alternative method involves immobilising the antigen in the biological sample and then exposing the immobilised antigen to specific antibody that may or may not be labelled with a reporter molecule. Depending on the amount of target and the strength of the reporter molecule signal, a bound antigen may be detectable by direct labelling with the antibody. Alternatively, a second labelled antibody, specific to the first antibody is exposed to the target-first antibody complex to form a target-first antibody-second antibody tertiary complex. The complex is detected by the signal emitted by the reporter molecule.

The association of the reporter molecule with the antigen-binding molecule may include the following: (a) direct attachment of the reporter molecule to the antigen-binding molecule; (b) indirect attachment of the reporter molecule to the antigen-binding molecule; *i.e.*, attachment of the reporter molecule to another assay reagent which subsequently binds to the antigen-binding molecule; and (c) attachment to a subsequent reaction product of the antigen-binding molecule.

The reporter molecule may be selected from a group including a chromogen, a catalyst, an enzyme, a fluorochrome, a chemiluminescent molecule, a lanthanide ion such as Europium (Eu^{34}), a radioisotope and a direct visual label.

In the case of a direct visual label, use may be made of a colloidal metallic or non-metallic particle, a dye particle, an enzyme or a substrate, an organic polymer, a latex particle, a liposome, or other vesicle containing a signal producing substance and the like.

A large number of enzymes suitable for use as reporter molecules is disclosed in United States Patent Specifications U.S. 4,366,241, U.S. 4,843,000, and U.S. 4,849,338. Suitable enzymes useful in the present invention include alkaline phosphatase, horseradish peroxidase, luciferase, β-galactosidase, glucose oxidase, lysozyme, malate dehydrogenase and the like. The enzymes may be used alone or in combination with a second enzyme that is in solution.

Suitable fluorochromes include, but are not limited to, fluorescein isothiocyanate (FITC), tetramethylrhodamine isothiocyanate (TRITC), R-Phycoerythrin (RPE), and Texas Red. Other exemplary fluorochromes include those discussed by Dower *et al.* (International Publication WO 93/06121). Reference also may be made to the fluorochromes described in U.S. Patents 5,573,909 (Singer *et al*), 5,326,692 (Brinkley *et al*). Alternatively, reference may be made to the fluorochromes described in U.S. Patent Nos. 5,227,487, 5,274,113, 5,405,975, 5,433,896, 5,442,045, 5,451,663, 5,453,517, 5,459,276, 5,516,864, 5,648,270 and 5,723,218.

In the case of an enzyme immunoassay, an enzyme is conjugated to the second antibody, generally by means of glutaraldehyde or periodates. As will be readily recognised, however, a wide variety of different conjugation techniques exist which are readily available to the skilled artisan. The substrates to be used with the specific enzymes are generally chosen for the production of, upon hydrolysis by the corresponding enzyme, a detectable colour change. Examples of suitable enzymes include those described *supra*. It is also possible to employ fluorogenic substrates, which yield a fluorescent product rather than the chromogenic substrates noted above. In all cases, the enzyme-labelled antibody is added to the first antibody-antigen complex. It is then allowed to bind, and excess reagent is washed away. A solution containing the appropriate substrate is then added to the complex of antibody-antigen-antibody. The substrate will react with the enzyme linked to the second antibody, giving a qualitative visual signal, which may be further quantitated,

usually spectrophotometrically, to give an indication of the amount of antigen which was present in the sample.

Alternately, fluorescent compounds, such as fluorescein, rhodamine and the lanthanide, europium (EU), may be chemically coupled to antibodies without altering their binding capacity. When activated by illumination with light of a particular wavelength, the fluorochrome-labelled antibody adsorbs the light energy, inducing a state to excitability in the molecule, followed by emission of the light at a characteristic colour visually detectable with a light microscope. The fluorescent-labelled antibody is allowed to bind to the first antibody-antigen complex. After washing off the unbound reagent, the remaining tertiary complex is then exposed to light of an appropriate wavelength. The fluorescence observed indicates the presence of the antigen of interest. Immunofluorometric assays (IFMA) are well established in the art. However, other reporter molecules, such as radioisotope, chemiluminescent or bioluminescent molecules may also be employed.

It will be well understood that other means of testing target polypeptide (*e.g.*, IGFBP) levels are available, including, for instance, those involving testing for an altered level of IGFBP binding activity to an insulin-like growth factor (IGF), or Western blot analysis of IGFBP protein levels in tissues, cells or fluids using anti-IGFBP antigen-binding molecule, or assaying the amount of antigen-binding molecule or other IGFBP binding partner which is not bound to a sample, and subtracting from the total amount of antigen-binding molecule or binding partner added.

3. Identification of target molecule modulators

The invention also provides agents that modulate the level and/or functional activity of an expression product of a gene selected from an *IGFBP* gene or a gene relating to the same regulatory or biosynthetic pathway as an *IGFBP* gene. The method includes contacting a preparation comprising target molecules selected from (1) all or part of an expression product of an *IGFBP* gene or a gene relating to the same regulatory or biosynthetic pathway as the *IGFBP* gene, or a variant or derivative of that expression product, or (2) a genetic sequence which regulates expression of the expression product mentioned in (1) and which is operably connected to a marker gene, with a test agent and detecting a change in the level and/or functional activity of all or part of that expression product or variant or derivative thereof or a change in the level of expression of the marker gene or a change in the phenotype of a cell.

In accordance with the present invention, agents that test positive for such interaction are useful for treating and/or modulating tumorigenesis and especially for treating or preventing a liver cancer. Assays of a suitable nature for detecting, measuring or otherwise determining modulation of tumorigenesis (*e.g.*, such as by detecting cell proliferation) are known to persons of skill in the art. For example, tumorigenesis-modulating agents could be tested for their ability to modulate cell proliferation. Typically, for cell proliferation, cell number is determined, directly, by microscopic or electronic enumeration, or indirectly, by the use of chromogenic dyes, incorporation of radioactive precursors or measurement of metabolic activity of cellular enzymes. An exemplary cell proliferation assay comprises culturing cells in the presence or absence of a test compound, and detecting cell proliferation by, for example, measuring incorporation of tritiated thymidine or by colorimetric assay based on the metabolic breakdown of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) (Mosman, 1983, *J. Immunol. Meth.* **65**: 55-63). Cancer or tumour markers are known for a variety of cell or tissue types. Cells or tissues expressing cancer or tumour markers may be detected using monoclonal antibodies, polyclonal antisera or other antigen-binding molecules that are immuno-interactive with these tumour markers or by using nucleic acid analysis techniques, including, for example, detecting the level or presence of tumour marker-encoding polynucleotides.

Modulators contemplated by the present invention includes agonists and antagonists of *IGFBP* gene expression include antisense molecules, ribozymes and co-suppression molecules. Agonists include molecules which increase promoter activity or interfere with negative mechanisms. Agonists of *IGFBP* include molecules which overcome any negative regulatory mechanism. Antagonists of IGFBP polypeptides include antibodies and inhibitor peptide fragments.

Candidate agents encompass numerous chemical classes, though typically they are organic molecules, preferably small organic compounds having a molecular weight of more than 50 and less than about 2,500 Dalton. Candidate agents comprise functional groups necessary for structural interaction with proteins, particularly hydrogen bonding, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups. The candidate agents often comprise cyclical carbon or heterocyclic structures and/or aromatic or polyaromatic structures substituted with one or more of the above functional groups. Candidate agents are also found among

biomolecules including, but not limited to: peptides, saccharides, fatty acids, steroids, purines, pyrimidines, derivatives, structural analogues or combinations thereof.

Small (non-peptide) molecule modulators of IGFBP–1, –2 and –3 are particularly preferred. In this regard, small molecules are particularly preferred because such molecules are more readily absorbed after oral administration, have fewer potential antigenic determinants, and/or are more likely to cross the cell membrane than larger, protein-based pharmaceuticals. Small organic molecules may also have the ability to gain entry into an appropriate cell and affect the expression of a gene (*e.g.*, by interacting with the regulatory region or transcription factors involved in gene expression); or affect the activity of a gene by inhibiting or enhancing the binding of accessory molecules.

Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural or synthetically produced libraries and compounds are readily modified through conventional chemical, physical and biochemical means, and may be used to produce combinatorial libraries. Known pharmacological agents may be subjected to directed or random chemical modifications, such as acylation, alkylation, esterification, amidification, etc. to produce structural analogues.

Screening may also be directed to known pharmacologically active compounds and chemical analogues thereof.

Screening for modulatory agents according to the invention can be achieved by any suitable method. For example, the method may include contacting a cell comprising a polynucleotide corresponding to an *IGFBP* gene or a gene belonging to the same regulatory or biosynthetic pathway as an *IGFBP* gene, with an agent suspected of having said modulatory activity and screening for the modulation of the level and/or functional activity of a protein encoded by said polynucleotide, or the modulation of the level of a transcript encoded by the polynucleotide, or the modulation of the activity or expression of a downstream cellular target of said protein or said transcript. Detecting such modulation can be achieved utilising techniques including, but not restricted to, ELISA, cell-based ELISA, filter-binding ELISA, inhibition ELISA, Western blots, immunoprecipitation, slot or dot blot assays, immunostaining, RIA, scintillation proximity assays, fluorescent immunoassays using antigen-binding molecule conjugates or antigen conjugates of fluorescent substances such as fluorescein or rhodamine, Ouchterlony double diffusion analysis, immunoassays employing an avidin-biotin or a streptavidin-biotin detection

system, and nucleic acid detection assays including reverse transcriptase polymerase chain reaction (RT-PCR).

It will be understood that a polynucleotide from which a target molecule of interest is regulated or expressed may be naturally occurring in the cell which is the subject of testing or it may have been introduced into the host cell for the purpose of testing. Further, the naturally-occurring or introduced polynucleotide may be constitutively expressed – thereby providing a model useful in screening for agents which downregulate expression of an encoded product of the sequence wherein said downregulation can be at the nucleic acid or protein level – or may require activation – thereby providing a model useful in screening for agents that upregulate expression of an encoded product of the sequence. Further, to the extent that a polynucleotide is introduced into a cell, that polynucleotide may comprise the entire coding sequence which codes for a target protein or it may comprise a portion of that coding sequence (*e.g.* a domain such as an IGF-binding domain) or a portion that regulates expression of a product encoded by the polynucleotide (*e.g.*, a promoter). For example, the promoter that is naturally associated with the polynucleotide may be introduced into the cell that is the subject of testing. In this regard, where only the promoter is utilised, detecting modulation of the promoter activity can be achieved, for example, by operably linking the promoter to a polynucleotide encoding a suitable marker including, but not restricted to, green fluorescent protein (GFP), luciferase, ß-galactosidase and catecholamine acetyl transferase (CAT). Modulation of expression may be determined by measuring the activity associated with the marker-encoding polynucleotide.

In another example, the subject of detection could be a downstream regulatory target of the target molecule, rather than target molecule itself or a marker-encoding polynucleotide operably linked to a promoter of a gene encoding a product the expression of which is regulated by the target protein.

These methods provide a mechanism for performing high throughput screening of putative modulatory agents such as proteinaceous or non-proteinaceous agents comprising synthetic, combinatorial, chemical and natural libraries. These methods will also facilitate the detection of agents which bind either the polynucleotide encoding the target molecule or which modulate the expression of an upstream molecule, which subsequently modulates the expression of the polynucleotide encoding the target molecule. Accordingly, these

methods provide a mechanism of detecting agents that either directly or indirectly modulate the expression and/or activity of a target molecule according to the invention.

In a series of preferred embodiments, the present invention provides assays for identifying small molecules or other compounds (*i.e.*, modulatory agents) which are capable of inducing or inhibiting the level and/or or functional activity of target molecules according to the invention. The assays may be performed *in vitro* using non-transformed cells, immortalised cell lines, or recombinant cell lines. In addition, the assays may detect the presence of increased or decreased expression of genes or production of proteins on the basis of increased or decreased mRNA expression (using, for example, nucleic acid probes corresponding or complementary to a transcript of an *IGFBP* gene or of a gene relating to the same regulatory or biosynthetic pathway as the *IGFBP* gene), increased or decreased levels of protein products (using, for example, antigen binding molecules that are immuno-interactive with a polypeptide encoded by the aforementioned genes), or increased or decreased levels of expression of a marker gene (*e.g.*, GFP, β-galactosidase or luciferase) operatively linked to a target molecule-related gene regulatory region in a recombinant construct.

Thus, for example, one may culture cells which produce a particular target molecule and add to the culture medium one or more test compounds. After allowing a sufficient period of time (*e.g.*, 6-72 hours) for the compound to induce or inhibit the level and/or functional activity of the target molecule, any change in said level from an established baseline may be detected using any of the techniques described above and well known in the art. In particularly preferred embodiments, the cells are epithelial cells. Using suitable nucleic acid probes and/or antigen-binding molecules as for example described above, detection of changes in the level and or functional activity of a target molecule, and thus identification of the compound as agonist or antagonist of the target molecule, requires only routine experimentation.

In particularly preferred embodiments, a recombinant assay is employed in which a marker gene encoding, for example, GFP, β-galactosidase or luciferase is operably linked to the 5' regulatory regions of a target molecule related gene. Such regulatory regions may be easily isolated and cloned by one of ordinary skill in the art in light of the present disclosure. The marker gene and regulatory regions are joined in-frame (or in each of the three possible reading frames) so that transcription and translation of the marker gene may proceed under the control of the regulatory elements of the target molecule related gene.

The recombinant construct may then be introduced into any appropriate cell type although mammalian cells are preferred, and human cells are most preferred. The transformed cells may be grown in culture and, after establishing the baseline level of expression of the marker gene, test compounds may be added to the medium. The ease of detection of the expression of the marker gene provides for a rapid, high throughput assay for the identification of agonists or antagonists of the target molecules of the invention.

Compounds identified by this method will have potential utility in modifying the expression of target molecule related genes *in vivo*. These compounds may be further tested in the animal models to identify those compounds having the most potent *in vivo* effects. In addition, as described above with respect to small molecules having target polypeptide binding activity, these molecules may serve as "lead compounds" for the further development of pharmaceuticals by, for example, subjecting the compounds to sequential modifications, molecular modelling, and other routine procedures employed in rational drug design.

In another embodiment, a method of identifying agents that inhibit IGFBP activity is provided in which a purified preparation of an IGFBP protein in the presence and absence of a candidate agent under conditions in which the IGFBP is active, and the level of IGFBP activity is measured by a suitable assay. For example, an IGFBP inhibitor can be identified by measuring the ability of a candidate agent to decrease IGFBP activity in a cell (*e.g.*, a liver cell). In this method, a cell that is capable of expressing an *IGFBP* is exposed to, or cultured in the presence and absence of, the candidate agent under conditions in which the IGFBP is active in the cell, and an activity such as tumorigenesis is detected. An agent tests positive if it inhibits any of these activities.

In yet another embodiment, random peptide libraries consisting of all possible combinations of amino acids attached to a solid phase support may be used to identify peptides that are able to bind to a target molecule or to a functional domain thereof. Identification of molecules that are able to bind to a target molecule may be accomplished by screening a peptide library with a recombinant soluble target molecule. The target molecule may be purified, recombinantly expressed or synthesised by any suitable technique. Such molecules may be conveniently prepared by a person skilled in the art using standard protocols as for example described in Sambrook, *et al.*, (1989, *supra*) in particular Sections 16 and 17; Ausubel *et al.*, (1994-1998, *supra*), in particular Chapters 10 and 16; and Coligan *et al.*, (1995-1997, *supra*), in particular Chapters 1, 5 and 6.

Alternatively, a target polypeptide according to the invention may be synthesised using solution synthesis or solid phase synthesis as described, for example, in Chapter 9 of Atherton and Shephard (*supra*) and in Roberge *et al* (1995, *Science* **269**: 202).

To identify and isolate the peptide/solid phase support that interacts and forms a complex with a target molecule, preferably a target polypeptide, it may be necessary to label or "tag" the target polypeptide. The target polypeptide may be conjugated to any suitable reporter molecule, including enzymes such as alkaline phosphatase and horseradish peroxidase and fluorescent reporter molecules such as fluorescein isothyiocynate (FITC), phycoerythrin (PE) and rhodamine. Conjugation of any given reporter molecule, with target polypeptide, may be performed using techniques that are routine in the art. Alternatively, target polypeptide expression vectors may be engineered to express a chimeric target polypeptide containing an epitope for which a commercially available antigen-binding molecule exists. The epitope specific antigen-binding molecule may be tagged using methods well known in the art including labelling with enzymes, fluorescent dyes or coloured or magnetic beads.

For example, the "tagged" target polypeptide conjugate is incubated with the random peptide library for 30 minutes to one hour at 22° C to allow complex formation between target polypeptide and peptide species within the library. The library is then washed to remove any unbound target polypeptide. If the target polypeptide has been conjugated to alkaline phosphatase or horseradish peroxidase the whole library is poured into a petri dish containing a substrate for either alkaline phosphatase or peroxidase, for example, 5-bromo-4-chloro-3-indoyl phosphate (BCIP) or 3,3',4,4"-diamnobenzidine (DAB), respectively. After incubating for several minutes, the peptide/solid phase-target polypeptide complex changes colour, and can be easily identified and isolated physically under a dissecting microscope with a micromanipulator. If a fluorescently tagged target polypeptide has been used, complexes may be isolated by fluorescent activated sorting. If a chimeric target polypeptide having a heterologous epitope has been used, detection of the peptide/target polypeptide complex may be accomplished by using a labelled epitope specific antigen-binding molecule. Once isolated, the identity of the peptide attached to the solid phase support may be determined by peptide sequencing.

4. Method of modulating IGFBP activity

The invention also extends to a method for modulating liver cell proliferation or tumorigenesis, comprising contacting said cell with an agent for a time and under

conditions sufficient to modulate the level and/or functional activity of an *IGFBP* expression product, especially an IGFBP polypeptide.

In the following embodiments, an agent is suitably used to reduce, repress or otherwise inhibit cell proliferation or tumorigenesis. In one embodiment, the agent increases the level and/or functional activity of IGFBP-3. Any suitable *IGFBP-3* inducers or stabilising/activating agents may be used in this regard and these can be identified or produced by methods for example disclosed in Section 3. In an alternate embodiment, the agent decreases the level and/or functional activity of IGFBP-1 or IGFBP-2. Suitable IGFBP-1 or IGFBP-2 inhibitors may be identified or produced by methods for example disclosed in Section 3. For example, a suitable IGFBP inhibitor may comprise oligoribonucleotide sequences, that include anti-sense RNA and DNA molecules and ribozymes that function to inhibit the translation of IGFBP protein-encoding mRNA.

Anti-sense RNA and DNA molecules act to directly block the translation of mRNA by binding to targeted mRNA and preventing protein translation. In regard to antisense DNA, oligodeoxyribonucleotides derived from the translation initiation site, *e.g.*, between -10 and +10 regions of a gene encoding a polypeptide according to the invention, are preferred. Ribozymes are enzymatic RNA molecules capable of catalysing the specific cleavage of RNA. The mechanism of ribozyme action involves sequence specific hybridisation of the ribozyme molecule to complementary target RNA, followed by a endonucleolytic cleavage. Within the scope of the invention are engineered hammerhead motif ribozyme molecules that specifically and efficiently catalyse endonucleolytic cleavage of *IGFBP*, especially *IGFBP–1* or *-2* RNA sequences. Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, which include the following sequences, GUA, GUU and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides corresponding to the region of the target gene containing the cleavage site may be evaluated for predicted structural features such as secondary structure that may render the oligonucleotide sequence unsuitable. The suitability of candidate targets may also be evaluated by testing their accessibility to hybridisation with complementary oligonucleotides, using ribonuclease protection assays.

Both anti-sense RNA and DNA molecules and ribozymes may be prepared by any method known in the art for the synthesis of RNA molecules. These include techniques for chemically synthesising oligodeoxyribonucleotides well known in the art such as for

example solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by *in vitro* and *in vivo* transcription of DNA sequences encoding the antisense RNA molecule. Such DNA sequences may be incorporated into a wide variety of vectors which incorporate suitable RNA polymerase promoters such as the T7 or SP6 polymerase promoters. Alternatively, antisense cDNA constructs that synthesise antisense RNA constitutively or inducibly, depending on the promoter used, can be introduced stably into cell lines.

Various modifications to the DNA molecules may be introduced as a means of increasing intracellular stability and half-life. Possible modifications include but are not limited to the addition of flanking sequences of ribo- or deoxy- nucleotides to the 5' and/or 3' ends of the molecule or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the oligodeoxyribonucleotide backbone.

5. Methods of preparing an IGFBP polypeptide

An IGFBP polypeptide or variant thereof may be prepared by any suitable procedure known to those of skill in the art. For example, the polypeptides may be prepared by a procedure including the steps of (a) preparing a recombinant polynucleotide comprising a nucleotide sequence encoding a polypeptide comprising the sequence set forth in any one of SEQ ID NO: 2, 4 or 6, or variant or derivative of these, which nucleotide sequence is operably linked to transcriptional and translational regulatory nucleic acid; (b) introducing the recombinant polynucleotide into a suitable host cell; (c) culturing the host cell to express recombinant polypeptide from said recombinant polynucleotide; and (d) isolating the recombinant polypeptide. Preferred nucleotide sequences include, but are not limited to the sequences set forth in SEQ ID NO: 1, 3 and 5.

The recombinant polynucleotide is preferably in the form of an expression vector that may be a self-replicating extra-chromosomal vector such as a plasmid, or a vector that integrates into a host genome. The transcriptional and translational regulatory nucleic acid will generally be appropriate for the host cell used for expression. Numerous types of appropriate expression vectors and suitable regulatory sequences are known in the art for a variety of host cells. Typically, the transcriptional and translational regulatory nucleic acid may include, but is not limited to, promoter sequences, leader or signal sequences, ribosomal binding sites, transcriptional start and stop sequences, translational start and termination sequences, and enhancer or activator sequences. Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be

either naturally occurring promoters, or hybrid promoters that combine elements of more than one promoter.

In a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selection genes are well known in the art and will vary with the host cell used.

The expression vector may also include a fusion partner (typically provided by the expression vector) so that the recombinant polypeptide of the invention is expressed as a fusion polypeptide with said fusion partner. The main advantage of fusion partners is that they assist identification and/or purification of said fusion polypeptide. In order to express said fusion polypeptide, it is necessary to ligate a polynucleotide according to the invention into the expression vector so that the translational reading frames of the fusion partner and the polynucleotide coincide. Well known examples of fusion partners include, but are not limited to, glutathione-S-transferase (GST), Fc potion of human IgG, maltose binding protein (MBP) and hexahistidine (HIS_6), which are particularly useful for isolation of the fusion polypeptide by affinity chromatography. For the purposes of fusion polypeptide purification by affinity chromatography, relevant matrices for affinity chromatography are glutathione-, amylose-, and nickel- or cobalt-conjugated resins respectively. Many such matrices are available in "kit" form, such as the QIAexpress™ system (Qiagen) useful with (HIS_6) fusion partners and the Pharmacia GST purification system. In a preferred embodiment, the recombinant polynucleotide is expressed in the commercial vector pFLAG as described more fully hereinafter.

Another fusion partner well known in the art is green fluorescent protein (GFP). This fusion partner serves as a fluorescent "tag" which allows the fusion polypeptide of the invention to be identified by fluorescence microscopy or by flow cytometry. The GFP tag is useful when assessing subcellular localisation of the fusion polypeptide of the invention, or for isolating cells which express the fusion polypeptide of the invention. Flow cytometric methods such as fluorescence activated cell sorting (FACS) are particularly useful in this latter application.

Preferably, the fusion partners also have protease cleavage sites, such as for Factor X_a or Thrombin, which allow the relevant protease to partially digest the fusion polypeptide of the invention and thereby liberate the recombinant polypeptide of the invention therefrom. The liberated polypeptide can then be isolated from the fusion partner by subsequent chromatographic separation.

Fusion partners according to the invention also include within their scope "epitope tags", which are usually short peptide sequences for which a specific antibody is available. Well known examples of epitope tags for which specific monoclonal antibodies are readily available include c-Myc, influenza virus, haemagglutinin and FLAG tags.

The step of introducing into the host cell the recombinant polynucleotide may be effected by any suitable method including transfection, and transformation, the choice of which will be dependent on the host cell employed. Such methods are well known to those of skill in the art.

Recombinant polypeptides of the invention may be produced by culturing a host cell transformed with an expression vector containing nucleic acid encoding a polypeptide, biologically active fragment, variant or derivative according to the invention. The conditions appropriate for protein expression will vary with the choice of expression vector and the host cell. This is easily ascertained by one skilled in the art through routine experimentation.

Suitable host cells for expression may be prokaryotic or eukaryotic. One preferred host cell for expression of a polypeptide according to the invention is a bacterium. The bacterium used may be *Escherichia coli*. Alternatively, the host cell may be an insect cell such as, for example, *SF9* cells that may be utilised with a baculovirus expression system.

The recombinant protein may be conveniently prepared by a person skilled in the art using standard protocols as for example described in Sambrook, *et al.*, 1989, in particular Sections 16 and 17; Ausubel *et al.*, (1994-1998), in particular Chapters 10 and 16; and Coligan *et al.*, (1995-1997), in particular Chapters 1, 5 and 6.

Alternatively, the polypeptide, fragments, variants or derivatives of the invention may be synthesised using solution synthesis or solid phase synthesis as described, for example, in Chapter 9 of Atherton and Shephard (*supra*) and in Roberge *et al* (1995).

6. Compositions

The modulatory agents described in Sections 3 and 4 (therapeutic agents) can be used as actives for modulating liver cell proliferation or tumorigenesis or for the treatment or prophylaxis of a liver cancer, particularly hepatocellular carcinoma, or related conditions as, for example, described below. These therapeutic agents can be administered to a patient either by themselves, or in pharmaceutical compositions where they are mixed with a suitable pharmaceutically acceptable carrier.

Accordingly, the invention also provides a composition for treatment and/or prophylaxis of a liver cancer or related condition, comprising an IGFB-3 polypeptide and optionally a pharmaceutically acceptable carrier. The invention also provides a composition for treatment and/or prophylaxis of a liver cancer or related condition, comprising an a modulatory agent that modulates the level and/or functional activity of an IGFBP including, but not limited to, IGFBP–1, IGFBP–2 and IGFBP–3, and optionally a pharmaceutically acceptable carrier.

In one embodiment, the composition comprises a modulatory agent that reduces the level and/or functional activity of IGFBP–1 and/or IGFBP–2, together with a pharmaceutically acceptable carrier. In another embodiment, the composition comprises a modulatory agent that enhances the level and/or functional activity of IGFBP–3, together with a pharmaceutically acceptable carrier.

Depending on the specific conditions being treated, therapeutic agents may be formulated and administered systemically or locally. Techniques for formulation and administration may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, Pa., latest edition. Suitable routes may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intramedullary injections, as well as intrathecal, direct intraventricular, intravenous, intraperitoneal, intranasal, or intraocular injections. For injection, the therapeutic agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks' solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art. Intra-muscular and subcutaneous injection is appropriate, for example, for administration of immunogenic compositions, vaccines and DNA vaccines.

The agents can be formulated readily using pharmaceutically acceptable carriers well known in the art into dosages suitable for oral administration. Such carriers enable the compounds of the invention to be formulated in dosage forms such as tablets, pills, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. These carriers may be selected from sugars, starches, cellulose and its derivatives, malt, gelatine, talc, calcium sulphate, vegetable oils, synthetic oils, polyols, alginic acid, phosphate buffered solutions, emulsifiers, isotonic saline, and pyrogen-free water.

Pharmaceutical compositions suitable for use in the present invention include compositions wherein the active ingredients are contained in an effective amount to achieve its intended purpose. The dose of agent administered to a patient should be sufficient to effect a beneficial response in the patient over time such as a reduction in the symptoms associated with a liver cancer or related condition. The quantity of the agent(s) to be administered may depend on the subject to be treated inclusive of the age, sex, weight and general health condition thereof. In this regard, precise amounts of the agent(s) for administration will depend on the judgement of the practitioner. In determining the effective amount of the agent to be administered in the treatment or prophylaxis of the liver cancer or related condition, the physician may evaluate tissue levels of an IGFBP polypeptide, and progression of the disorder. In any event, those of skill in the art may readily determine suitable dosages of the therapeutic agents of the invention.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilisers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Pharmaceutical preparations for oral use can be obtained by combining the active compounds with solid excipient, optionally grinding a resulting mixture, and processing the mixture of granules, after adding suitable auxiliaries, if desired, to obtain tablets or dragee cores. Suitable excipients are, in particular, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as., for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl-cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate. Such compositions may be prepared by any of the methods of pharmacy but all methods include the step of bringing into association one or more therapeutic agents as described above with the carrier which constitutes one or more necessary ingredients. In

general, the pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilising processes.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterise different combinations of active compound doses.

Pharmaceutical which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticiser, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilisers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilisers may be added.

Dosage forms of the therapeutic agents of the invention may also include injecting or implanting controlled releasing devices designed specifically for this purpose or other forms of implants modified to act additionally in this fashion. Controlled release of an agent of the invention may be effected by coating the same, for example, with hydrophobic polymers including acrylic resins, waxes, higher aliphatic alcohols, polylactic and polyglycolic acids and certain cellulose derivatives such as hydroxypropylmethyl cellulose. In addition, controlled release may be effected by using other polymer matrices, liposomes and/or microspheres.

Therapeutic agents of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but not limited to hydrochloric, sulphuric, acetic, lactic, tartaric, malic, succinic, *etc*. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms.

For any compound used in the method of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes

the IC50 as determined in cell culture (*e.g.*, the concentration of a test agent, which achieves a half-maximal inhibition or enhancement of IGFBP activity). Such information can be used to more accurately determine useful doses in humans.

Toxicity and therapeutic efficacy of such therapeutic agents can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD50 (the dose lethal to 50% of the population) and the ED50 (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio LD50/ED50. Compounds that exhibit large therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED50 with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilised. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See for example Fingl *et al.*, 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p1).

Dosage amount and interval may be adjusted individually to provide plasma levels of the active agent which are sufficient to maintain IGFBP-inhibitory or enhancement effects. Usual patient dosages for systemic administration range from 1-2000 mg/day, commonly from 1-250 mg/day, and typically from 10-150 mg/day. Stated in terms of patient body weight, usual dosages range from 0.02-25 mg/kg/day, commonly from 0.02-3 mg/kg/day, typically from 0.2-1.5 mg/kg/day. Stated in terms of patient body surface areas, usual dosages range from 0.5-1200 mg/m^2/day, commonly from 0.5-150 mg/m^2/day, typically from 5-100 mg/m^2/day.

Alternately, one may administer the compound in a local rather than systemic manner, for example, *via* injection of the compound directly into a tissue, which is preferably a heart muscle tissue or a liver tissue, often in a depot or sustained release formulation. Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with tissue-specific antibody. The liposomes will be targeted to and taken up selectively by the tissue. In cases of local administration or selective uptake, the effective local concentration of the agent may not be related to plasma concentration.

Thus, the present invention also contemplates a method of gene therapy of a mammal. Such a method utilises a gene therapy construct which includes an isolated polynucleotide comprising a nucleotide sequence encoding an IGFBP–3 or a biologically active fragment thereof, wherein said polynucleotide is ligated into a gene therapy vector which provides one or more regulatory sequences that direct expression of said polynucleotide in said mammal.

Typically, gene therapy vectors are derived from viral DNA sequences such as adenovirus, adeno-associated viruses, herpes-simplex viruses and retroviruses. Suitable gene therapy vectors currently available to the skilled person may be found, for example, in Robbins *et al.*, 1998.

If "anti-sense" therapy is contemplated (*e.g.*, IGFBP–1 and/or –2), then one or more selected portions of a *IGFBP-1* and/or *IGFBP-2* nucleic acid may be oriented 3'→ 5' in the gene therapy vector.

Administration of the gene therapy construct to said mammal, preferably a human, may include delivery *via* direct oral intake, systemic injection, or delivery to selected tissue(s) or cells, or indirectly *via* delivery to cells isolated from the mammal or a compatible donor. An example of the latter approach would be stem-cell therapy, wherein isolated stem cells having potential for growth and differentiation are transfected with the vector comprising an *IGFBP* nucleic acid. The stem-cells are cultured for a period and then transferred to the mammal being treated.

Delivery of said gene therapy construct to cells or tissues of said mammal or said compatible donor may be facilitated by microprojectile bombardment, liposome mediated transfection (*e.g.*, lipofectin or lipofectamine), electroporation, calcium phosphate or DEAE-dextran-mediated transfection, for example. A discussion of suitable delivery methods may be found in Chapter 9 of Ausubel *et al.*, (1994-1998, *supra*). For example, a nucleic acid encoding IGFBP–3 may be introduced into a cell to enhance the ability of that cell to differentiate or to decrease proliferation. Alternatively, this may be achieved by introducing into a cell *IGFBP–1* and/or *IGFBP–2* antisense sequences such as 3'→ 5' oligonucleotides.

In an alternate embodiment, a polynucleotide encoding a modulatory agent of the invention may be used as a therapeutic or prophylactic composition in the form of a "naked DNA" composition as is known in the art. For example, an expression vector comprising said polynucleotide operably linked to a regulatory polynucleotide (*e.g.* a promoter,

transcriptional terminator, enhancer *etc*) may be introduced into an animal, preferably a mammal, where it causes production of a modulatory agent *in vivo*, preferably in a heart muscle tissue or a liver tissue.

The step of introducing the expression vector into a target cell or tissue will differ depending on the intended use and species, and can involve one or more of non-viral and viral vectors, cationic liposomes, retroviruses, and adenoviruses such as, for example, described in Mulligan, R.C., (1993). Such methods can include, for example:

A. Local application of the expression vector by injection (Wolff *et al.*, 1990), surgical implantation, instillation or any other means. This method can also be used in combination with local application by injection, surgical implantation, instillation or any other means, of cells responsive to the protein encoded by the expression vector so as to increase the effectiveness of that treatment. This method can also be used in combination with local application by injection, surgical implantation, instillation or any other means, of another factor or factors required for the activity of said protein.

B. General systemic delivery by injection of DNA, (Calabretta *et al.*, 1993), or RNA, alone or in combination with liposomes (Zhu *et al.*, 1993), viral capsids or nanoparticles (Bertling *et al.*, 1991) or any other mediator of delivery. Improved targeting might be achieved by linking the polynucleotide/expression vector to a targeting molecule (the so-called "magic bullet" approach employing, for example, an antigen-binding molecule), or by local application by injection, surgical implantation or any other means, of another factor or factors required for the activity of the protein encoded by said expression vector, or of cells responsive to said protein.

C. Injection or implantation or delivery by any means, of cells that have been modified *ex vivo* by transfection (for example, in the presence of calcium phosphate: Chen *et al.*, 1987, or of cationic lipids and polyamines: Rose *et al.*, 1991), infection, injection, electroporation (Shigekawa *et al.*, 1988) or any other way so as to increase the expression of said polynucleotide in those cells. The modification can be mediated by plasmid, bacteriophage, cosmid, viral (such as adenoviral or retroviral; Mulligan, 1993; Miller, 1992; Salmons *et al.*, 1993) or other vectors, or other agents of modification such as liposomes (Zhu *et al.*, 1993), viral capsids or nanoparticles (Bertling *et al.*, 1991), or any other mediator of modification. The use of cells as a delivery vehicle for genes or gene products has been described by Barr *et al.*, 1991 and by Dhawan *et al.*,

1991. Treated cells can be delivered in combination with any nutrient, growth factor, matrix or other agent that will promote their survival in the treated subject.

In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting examples.

EXAMPLES

EXAMPLE 1

Reduction of IGFBP-3 expression in human hepatocellular carcinoma

Methods

Collection of human HCC and adjacent non-tumour liver specimens

Tissue samples were obtained intra-operatively from tumours and adjacent non-tumour livers during liver resection for HCC in 28 patients treated at the Singapore General Hospital. The samples were snap frozen in liquid nitrogen and stored at -80° C until analysis. A similar set of samples was fixed in 10% formalin and paraffin embedded. The diagnosis of HCC was confirmed histologically in all cases. Prior written informed consent was obtained from all patients and the study received ethics board apprc val at both institutions.

Immunolocalisation of IGFBP-3

Formalin- and paraffin-embedded sections were used for IGFBP-3, IGF-II, and IGF-I immunolocalisation. This was performed using rabbit anti-human IGFBP-3 (1:500) (Upstate Biotechnology, Lake Placid, NY), rabbit anti-human IGF-I (3 μg/mL) (Santa Cruz Biotechnology, Santa Cruz, CA) and rabbit anti-human IGF-II (3 μg/mL) (Austral Biologicals, San Ramon, CA) antibodies as described (14). Non-specific staining was evaluated for each specimen using either a similar concentration of IgG or by absorbing the primary with appropriate specific immunogen. The slides were evaluated and intensity of the staining was scored. Specific staining was semi-quantitated by assigning a score of 0 to +++ based on increasing green fluorescence intensity. The results shown in TABLE 1 represent the average score obtained from twice staining.

Cell culture

Human hepatoma HepG2 cells were obtained from ATCC and maintained as monolayer cultures in Modified Eagle's Media (MEM) (Gibco, Grand Island, NY) supplemented with 10% foetal calf serum (FCS) (Gibco, Grand Island, NY) growth medium. For proliferation study, confluent cultures of HepG2 cells were trypsinised and plated at 2.0 x 10^4 cells in 24-well plates with growth medium. After 48h, the cell monolayers were rinsed twice with MEM phenol-red serum free (PSF) media and incubated further in PSF medium for 24h. After 24h, various concentrations of human

recombinant IGFBP-3 (hrIGFBP-3) (Celtrix, CA), human recombinant IGF-I (hrIGF-I, GroPep, Adelaide, Australia) or 25 ng/mL hrIGF-I in conjunction with various concentrations of hrIGFBP-3 were added in triplicate in SPF media for 48 h. Cell number was determined as previously described (14).

To determine the effect of exogenous IGF-II and autocrine IGF-II production on HepG2 cell proliferation, HepG2 cells were plated and grown as described above and the cells subsequently treated with rabbit pre-immune serum (1:800), 25 ng/mL hrIGF-II (GroPep, Adelaide, Australia), rabbit anti-human IGF-II antibody (2 μg/mL) or with combination for 48 h. Cell number was determined as described above.

Western Blotting

To determine changes in the expression of IGFBP-3 and IGF-IR, snap-frozen HCC tumours and normal adjacent liver tissues were thawed and homogenised in lysis buffer (1 mM CaC12, 1 mM MgC12, 1% NP-40, 1 μg/mL leupeptin, 1 μg/mL aprotinin, 1 μM PMSF, and 100 μM $NaV0_4$). Proteins were subjected to Western blot analysis as described (15). Blots were incubated with 1:5000 anti-human IGFBP-3 (Upstate Biotechnology), 1 μg/mL anti-human IGF-IR (Santa Cruz, CA), 0.5 μg/mL anti α-tubulin (Santa Cruz, CA) antibodies and 1:7500 horseradish peroxidase-conjugated donkey anti-mouse or anti-rabbit secondary antibody (Pierce, Rockford, Illinois). The blots were then visualised with a chemiluminescent detection system as described by the manufacturer (ECL, Amersham).

Results

In the adult human, the liver is the main source of circulating IGF-I and IGFBP-3 (12, 13). Since local expression of IGF-I is important in autocrine and paracrine stimulation and the IGFBP-3 appears to decrease the mitogenic activity of free IGF-I, presumably by competing with type I IGF receptors for the ligand, the IGFBP-3 levels in the HCC samples were determined. Total tissue lysates from both HCC cancers and adjacent morphologically normal liver cells were subjected to Western blot analysis. The resulting blots were incubated with IGFBP-3 antibody. Figure 1 shows that IGFBP-3 protein (a complex spanning 38-42 kDa) was either absent or low in HCC compared to adjacent normal tissues. Unlike IGFBP-3 detected in the serum, which comprises two forms (42 and 38 kDa bands) of equal intensity, the 38 kDa IGFBP-3 was the predominant form in all normal liver tissue. This band was undetectable in 28.5% (8 of 28) of HCC

samples examined (TABLE 1). Although the remaining tumours (20 of 28; 71.5%) expressed IGFBP–3, the level was significantly lower than that observed in normal adjacent tissue ($p<0.01$). Densitometric scanning showed IGFBP–3 levels to be 4 to 100-fold higher in normal liver tissue than in HCC. Subsequent blotting with anti-α-tubulin antibody showed relatively equal amounts of total protein loaded per lane (Figure 1). Figure 1C also demonstrated that proteolysis of IGFBP–3 was unlikely to be responsible for the low levels or absence of IGFBP–3 from HCC samples because low molecular weight IGFBP–3 fragments were not clearly detected on Western blot analysis of the tissue samples. The IGFBP–3 fragments, however, were detected in serum samples.

To confirm the above observation, immunohistochemical staining of HCC and normal adjacent liver tissues was performed using anti-human IGFBP–3 antibody which recognised both intact and fragmented IGFBP–3 protein. IGFBP-3 protein was clearly localised to almost all cells in the adjacent normal tissue (Figure 2A). However, this signal was absent (Figure 2B) in HCC, indicating that IGFBP–3 expression was either inactivated or decreased in these neoplastic cells.

To determine if HCC also expressed IGF–I and IGF–II, HCC and adjacent normal liver tissues were stained with anti-IGF–I and anti-IGF–II antibodies, respectively. Although normal adjacent tissues stained positively for IGF–I (Figure 3A), the expression was significantly decreased in HCC (Figure 3B). Approximately 64% (18/28) and 36% (10/28) of HCC samples examined had moderate and low IGF–I expression, respectively (TABLE 1) which is consistent with a previous report (28). IGF–II staining (Figures 3C and 3D), in contrast, was approximately equal (68%; 19 of 28) or more intense (32%; 9 of 28) in HCC compared to normal liver tissue (TABLE 1). These observations are in agreement with previous reports demonstrating over-expression of IGF–II in HCC (9-11). Our data suggests that IGF autocrine and paracrine loops exist in HCC.

Since IGFs stimulate growth responses in liver cells by binding to the IGF–I receptor (IGF–IR), IGF–IR levels in HCC and normal adjacent tissue samples were examined. Approximate 39% of HCC tumours had significantly lower IGF–IR levels than non-neoplastic adjacent liver tissues ($p<0.01$) (Figure 4, TABLE 1).

To test the hypothesis that IGF–I and IGF–II play a role in mediating tumour cell growth, HepG2 cells were treated with various concentrations of human recombinant IGF-I or IGF–II for 48 h. Figure 5A shows dose-dependent growth stimulation by IGF–I. Similar

effects were observed when HepG2 cells were treated with IGF–II (data not shown). Four-fold increase in cell number was observed at the dose of 25 ng/mL IGF–I ($p<0.01$).

As IGFBP–3 protein was undetectable or lost in the majority of HCC samples (Figure 1) and IGFBP–3 has been shown to inhibit cancer cells in an IGF-independent pathway (14), the inventors attempted to demonstrate the antiproliferative action of IGFBP–3 on HepG2 cells. Approximate 40% and 52% inhibition of basal proliferation by 250 ng/mL and 500 ng/mL hrIGFBP–3, respectively ($p<0.01$) was obtained (Figure 5B). Concentrations of rhIGFBP–3 up to 1 μg/mL did not further inhibit proliferation.

To determine if the proliferative action of IGF–I can be attenuated in the presence of IGFBP–3, HepG2 cells were treated with hr1GFBP–3 in the presence and absence of IGF–I for 48 h. As shown in Figure 5C, treatment of HepG2 with 25 ng/mL IGF–I for 48h resulted in a 2.8-fold increase in cell number ($p<0.01$). IGF-induced HepG2 proliferation was significantly attenuated ($p<0.01$) in the presence of 250 ng/mL IGFBP–3 and completely abolished at the concentration of 500 ng/mL. This result suggests that IGFBP-3 attenuated IGF–I-induced HepG2 proliferation by reducing IGF–I bioavailability.

To test the hypothesis that autocrine production of IGF–II by liver cancer cells plays a role in mediating liver cancer cell growth, HepG2 cells, which has been shown to secrete IGF–II (8), were treated with anti-IGF–II, IGF–II or with a combination for 48 h. Figure 5D shows a 3-fold induction of basal proliferation by 25 ng/mL IGF–II ($p<0.01$). This induction was significantly attenuated by IGF–II antibody ($p<0.01$), whereas pre-immune serum was not significantly affected. IGF–II antibody alone caused a 35% reduction in cell number ($p<0.05$). These results suggest that the rapid proliferation of liver cancer cells *in vivo* and *in vitro* may at least, in part, be a consequence of autocrine stimulation mediated by IGF–II expression.

EXAMPLE 2

Insulin-like growth factor binding protein levels in serum of hepatocellular carcinoma patients: Potential Markers for HCC

Methods

Serum collection from human HCC, prostate cancer, BPH and healthy patients.

Prior written informed consent was obtained from all patients and the study received ethics board approval at both institutions. HCC serum samples (n=116) were

obtained from patients from intra-venous access line during surgery (n=17) and from venous blood samples obtained from patients during index outpatient visit and on every monthly outpatient visit (n=99) at the Singapore General Hospital. The diagnosis of HCC was confirmed histologically in all cases. Serum from non-HCC (n=45), prostate cancer (n=29) and benign prostate hyperplasia (n=36) were obtained from venous blood samples obtained from patients during outpatient visit at Urology Centre, Singapore General Hospital. The serum was separated from blood samples and were snap frozen in liquid nitrogen and stored at -80° C until analysis.

Western blot analysis

To determine changes in the expression of IGFBP–3, IGFBP–1, IGFBP–2 and alpha fetoprotein, 3.5 μL serum were subjected to Western blot analysis as described (14). Blots were incubated with rabbit anti-human IGFBP–3 (1:5000), rabbit anti-bovine IGFBP–2 (1:2500), rabbit anti-human IGFBP–1 (1:2000) and rabbit anti-human-alpha fetoprotein (1:200) antibodies and 1:7500 horseradish peroxidase-conjugated donkey anti-mouse or anti-rabbit secondary antibody (Pierce, Rockford, Illinois). The blots were then visualised with a chemiluminescent detection system as described by the manufacturer (ECL, Amersham). IGFBP–1, IGFBP–2 and IGFBP–3 were purchased from (Upstate Biotechnology).

Results

In the adult human, the liver is the main source of circulating, IGF–I, IGFBP–1 and IGFBP–3 (12, 13). Since local expression of IGF–I is important in autocrine and paracrine stimulation and the IGFBPs appear to regulate the mitogenic activity of free IGF–I, the IGFBP–3, IGFBP–2 and IGFBP–1 levels in the HCC serum samples were determined. Serum from HCC patients were subjected to Western blot analysis. The serums derived from prostate cancer, BPH and healthy patients served as controls.

While IGFBP–3 detected in the serum comprises two forms, 42 and 38 kDa, these forms were undetectable in 70% (81 of 116) of HCC serum samples examined. Although the remaining HCC serum samples had detectable IGFBP–3, the levels were significantly lower than that observed in control serum samples (Figure 6A-6P). Unlike IGFBP–3, IGFBP–2 and IGFBP–1 levels were low in control serum samples. Both IGFBP–2 and IGFBP–1 levels were increased in HCC serum but not others. IGFBP–3 levels were inversely correlated with HCC while IGFBP–1 and IGFBP–2 levels were positively

correlated with HCC. The elevation of IGFBP–1 was more closely associated with HCC conditioned than IGFBP–2. Since AFP levels have been shown to be useful in detecting hepatocellular carcinoma (HCC), the AFP levels were determined. As shown for example in Figure 6E, approximately 20% of HCC serum samples did not show any increase in AFP. Thus, the serum AFP was not always correlated with the presence of HCC.

The disclosure of every patent, patent application, and publication cited herein is hereby incorporated herein by reference in its entirety.

The citation of any reference herein should not be construed as an admission that such reference is available as "Prior Art" to the instant application

Throughout the specification the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. Those of skill in the art will therefore appreciate that, in light of the instant disclosure, various modifications and changes can be made in the particular embodiments exemplified without departing from the scope of the present invention. All such modifications and changes are intended to be included within the scope of the appended claims.

TABLES

TABLE 1

Expression of IGFs, IGFBP–3 and IGF–IR in normal adjacent liver tissues and HCC Tumours

Genes Examined	Normal adjacent Liver Tissues (n - 28)	HCC Tumours (n = 28)
Intense staining of IGF–I	+++ (100%)	++ 18/28 (64%) + 10/28 (36%)
Intense staining of IGF–II	+++ (100%)	+++ 19/28 (68%) ++++ 9/28 (32%)
Expression levels of IGFBP–3 levels: (compared with normal and adjacent tissue of the same patient)		
• Normal	28/28 (100%)	0/28 (0%)
• Low	0/28 (0%)	8/28 (28.5%)
• Undetectable	0/28 (0%)	20/28 (71.5%)
IGF–IR expression levels: (compared with normal and adjacent tissue of the same patient)		
• Normal	28/28 (100%)	17/28 (61%)
• Low	0/28 (0%)	11/28 (39%)

BIBLIOGRAPHY

1. Okuda, K., Ohtsuki, T., Obata, H., Tomimatsu, M., Okazaki, N., Hasegawa, H., Nakajima, Y., and Ohnishi, K. Natural history of hepatocellular carcinoma and prognosis in relation to treatment. Study of 850 patients. Cancer, 56: 918-928, 1985.

2. Lai, E. C., Fan, S. T., Lo, C. M., Chu, K. M., Liu, C. L., and Wong, J. Hepatic resection for hepatocellular carcinoma. An audit of 343 patients. Ann.Surg., 221: 291-298, 1995.

3. Takenaka, K., Kawahara, N., Yamamoto, K., Kajiyama, K., Maeda, T., Itasaka, H., Shirabe, K., Nishizaki, T., Yanaga, K., and Sugimachi, K. Results of 280 liver resections for hepatocellular carcinoma. Arch.Surg.,131: 71-76, 1996.

4. Huguet, C., Stipa, F., and Gavelli, A. Primary hepatocellular cancer: Western experience. *In* L.Blumgart (ed.), Surgery of the liver and biliary tract, pp. 1365-1369. Churchill Livingstone: London, 2000.

5. Lai, E. and Wong, J. Hepatocellular carcinoma: the Asian experience. *In* L. Blumgart (ed.), Surgery of the Liver and the biliary tract, pp. 1349-1363. London: Churchill Livingstone, 1994.

6. Chan, E., Chow PKH, Tai, B. C., Martin D, and Soo KC Neoadjuvant and adjuvant thereapy for operable hepatocellular carcinoma. Cochrane Database System Rev., 2: 2000.

7. Scharf, J. G., Schmidt-Sandte, W., Pahernik, S. A., Ramadori, G., Braulke, T., and Hartmann, H. Characterization of the insulin-like growth factor axis in a human hepatoma cell line (PLC). Carcinogenesis, 19: 2121-2128, 1998.

8. Tsai, T. F., Yauk, Y. K., Chou, C. K., Ting, L. P., Chang, C., Hu, C. P., Han, S. H., and Su, T. S. Evidence of autocrine regulation in human hepatoma cell lines. Biochem.Biophys.Res.Commun., 153: 39-45, 1988.

9. Su, T. S., Liu, W. Y., Han, S. H., Jansen, M., Yang-Fen, T. L., P'eng, F. K., and Chou, C. K. Transcripts of the insulin-like growth factors I and II in human hepatoma. Cancer Res., 49: 1773-1777, 1989.

10. Ng, I. O., Lee, J. M., Srivastava, G., and Ng, M. Expression of insulin-like growth factor II mRNA in hepatocellular carcinoma. J.Gastroenterol.Hepatol., 13: 152-157, 1998.

11. Cariani, E., Lasserre, C., Seurin, D., Hamelin, B., Kemeny, F., Franco, D., Czech, M. P., Ullrich, A., and Brechot, C. Differential expression of insulin-like growth factor II mRNA in human primary liver cancers, benign liver tumors, and liver cirrhosis. Cancer Res., 48: 6844-6849, 1988.

12. Khandwala, H. M., McCutcheon, I. E., Flyvbjerg, A., and Friend, K. E. The effects of insulin-like growth factors on tumorigenesis and neoplastic growth [In Process Citation]. Endocr.Rev., 21: 215-244, 2000.

13. Clemmons, D. R. Insulin-like growth factor binding proteins and their role in controlling IGF actions. Cytokine and Growth Factor Reviews, 8: 45-62, 1997.

14. Huynh, H., Larsson, C., Narod, S., and Pollak, M. Tumour suppressor activity of the gene encoding mammary-derived growth inhibitor. Cancer Res., 55: 2225-2231, 1995.

15. Arany, E., Afford, S., Strain, A. J., Winwood, P. J., Arthur, M. J., and Hill, D. J. Differential cellular synthesis of insulin-like growth factor binding protein-1 (IGFBP-1) and IGFBP-3 within human liver. J.C1in.Endocrinol.Metab, 79: 1871-1876, 1994.

16. Huynh, H., Nickerson, T., Yang, X., and Pollak, M. Regulation of insulin-like growth factor I receptor by the pure antiestrogen ICI 182,780. Clinical Cancer Res., 2: 2037-2042, 1996.

17. Huynh, H., Yang, X., and Pollak, M. Estradiol and antiestrogens regulate a growth inhibitory insulin- like growth factor binding protein 3 autocrine loop in human breast cancer cells. J.Biol.Chem., 271: 1016-1021, 1996.

18. Oh, Y., Mullers, H. L., Ng, L., and Rosenfeld, R. G. Transforming growth factor-b-induced cell growth inhibition in human breast cancer cells is mediated through insulin-like growth factor-binding protein-3 action. J.Biol.Chem., 270: 13589-13592, 1995.

19. LeRoith, D., Adamo, M. L., Shemer, J., Lanau, F., Shen-Orr, Z., Yaron, A., Roberts, C. T., Clemmons, D. R., Sheikh, M. S., Shao, Z. M., and. Retinoic acid inhibits growth of breast cancer cell lines: the role of insulin-like growth factor binding proteins. Growth Regul., 3: 78-80, 1993.

20. Herman, J. G., Latif, F., Weng, Y., Lerman, M. I., Zbar, B., Liu, S., Samid, D., and et al. Silencing of the VHL tumor-suppressor gene by DNA methylation in renal carcinoma. Proc.Natl.Acad.Sci.USA, 91: 9700-9704, 1994.

21. Huynh, H., Alpert, L., and Pollak, M. Silencing of the mammary derived growth inhibitor (MDGI) gene in breast neoplasms is associated with epigenetic changes. Cancer Res., 56: 4865-4870, 1996.

22. Wong, I. H., Johnson, P. J., Lai, P. B., Lau, W. Y., and Lo, Y. M. Tumor-derived epigenetic changes in the plasma and serum of liver cancer patients. Implications for cancer detection and monitoring. Ann N.Y.Acad Sci., 906: 102-105, 2000.

23. Kamb, A., Gruis, N. A., Weaver-Feldhaus, J., Liu, Q., Harshman, K., Tavtigian, S. V., Stockert, E., Day, R. S., III, Johnson, B. E., and Skolnick, M. H. A cell cycle regulator potentially involved in genesis of many tumor types. Science, 264: 436-440, 1994.

24. Montesano, R., Hainaut, P., and Wild, C. P. Hepatocellular carcinoma: from gene to public health. J.Natl.Cancer Inst., 89: 1844-1851, 1997.

25. Chen, C. J., Yu, M. W., and Liaw, Y. F. Epidemiological characteristics and risk factors of hepatocellular carcinoma. J.Gastroenterol.Hepatol., 12: 5294-5308, 1997.

26. Oh, Y., Muller, H. L., Lamson, G., and Rosenfeld, R. G. Insulin-like growth factor (IGF)independent action of IGF-binding protein-3 in Hs578T human breast cancer cells. J.Biol.Chem., 268: 14964-14971, 1993.

27. Oh, Y., Muller, H. L., Pham, H., and Rosenfeld, R. G. Demonstration of receptors for insulin-like growth factor binding protein-3 on Hs578T human breast cancer cells. J.Biol.Chem., 268: 26045-26048, 1993.

28. Wakeling, A. E., Dukes, M., and Bowler, J. A potent specific pure antiestrogen with clinical potential. Cancer Res., Sl: 3867-3873, 1991.

29. Cameron, R. G., Imaida, K., Tsuda, H., and Ito, N. Promotive effects of steroids and bile acids on hepatocarcinogenesis initiated by diethylnitrosamine. Cancer Res., 42: 2426-2428, 1982.

30. Coe, J. E., Ishak, K. G., and Ross, M. J. Estrogen induction of hepatocellular carcinomas in Armenian hamsters. Hepatology, 11: 570-577, 1990.

CLAIMS

1. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient aberrant expression of a gene encoding an insulin-like growth factor binding protein (IGFBP).

2. The method of claim 1, wherein said IGFBP is selected from IGFBP–1, IGFBP–2 or IGFBP–3.

3. The method of claim 1, wherein said aberrant expression is detected by detecting a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said expression product.

4. The method of claim 1, wherein said aberrant expression is detected by detecting a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said expression product.

5. The method of claim 1, wherein said aberrant expression is detected by detecting a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said expression product.

6. The method of any one of claims 3 to 5, wherein the level and/or functional activity of said expression product in the biological sample obtained from said patient differs by at least 10% compared to the level and/or functional activity of said expression product in a biological sample obtained from normal patients or from patients who are not afflicted with said liver cancer.

7. The method of any one of claims 3 to 5, wherein said expression product is a transcript.

8. The method of any one of claims 3 to 5, wherein said expression product is a polypeptide.

9. The method of claim 1, wherein said liver cancer is liver cell carcinoma.

10. The method of claim 1, wherein said biological sample comprises a fluid selected from whole blood, plasma, lymph, saliva, mucus secretion and respiratory fluid.

11. The method of claim 1, wherein said biological sample comprises a circulatory fluid selected from serum or whole blood or a fractionated portion thereof.

12. The method of claim 1, wherein said biological sample comprises serum or a fractionated portion thereof.

13. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and/or

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and/or

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

14. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product.

15. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

16. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

17. A method for detecting the presence or diagnosing the risk of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient:

– a level and/or functional activity of an expression product of an *IGFBP–1* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–1* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–2* gene, which is elevated relative to a normal reference level and/or functional activity of said *IGFBP–2* expression product; and

– a level and/or functional activity of an expression product of an *IGFBP–3* gene, which is reduced relative to a normal reference level and/or functional activity of said *IGFBP–3* expression product.

18. A method for prognostic assessment of a liver cancer in a patient, comprising detecting in a biological sample obtained from said patient aberrant expression of at least one gene encoding an IGFBP.

19. Use of an agent in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of a gene selected from an *IGFBP* gene or a gene relating to the same regulatory or biosynthetic pathway as the *IGFBP* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a change in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a change in the level of expression of said marker gene.

20. Use of an agent that reduces the level and/or functional activity of an expression product of an *IGFBP–1* gene in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–1* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

21. Use of an agent that reduces the level and/or functional activity of an expression product of an *IGFBP–2* gene in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–2* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

22. Use of an agent that increases the level and/or functional activity of an expression product of an *IGFBP–3* gene in the manufacture of a medicament for treating and/or preventing a liver cancer, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–3* gene, or a variant or derivative of said expression product, or comprising a

genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting an increase in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or an increase in the level and/or functional activity of an expression product of said marker gene.

23. The use of claim 20 or claim 21, wherein said agent is an antisense oligonucleotide or ribozyme that binds to, or otherwise interacts specifically with, a polynucleotide selected from an *IGFBP–1* polynucleotide, an *IGFBP–2* polynucleotide or complement of these.

24. The use of claim 20 or claim 21, wherein said agent is an antigen-binding molecule that is immuno-interactive with IGFBP–1 and/or IGFBP–2.

25. The use of claim 22, wherein said agent comprises an IGFBP–3 polypeptide or a biologically active fragment thereof, or variant or derivative of these.

26. The use of claim 22, wherein said agent comprises a polynucleotide encoding an IGFBP–3 polypeptide or a biologically active fragment thereof, or variant or derivative of these, wherein said polynucleotide is operably connected to a regulatory polynucleotide.

27. Use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises at least a portion of an *IGFBP–3* polynucleotide operably linked, in the sense direction, to a regulatory polynucleotide.

28. Use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises an isolated *IGFBP–1* polynucleotide operably linked, in the antisense direction, to a regulatory polynucleotide.

29. Use of a vector in the manufacture of a medicament, which is optionally formulated with a pharmaceutically acceptable carrier, for the treatment and/or prophylaxis of a liver cancer or related condition, wherein said vector comprises an isolated *IGFBP–2*

polynucleotide operably linked, in the antisense direction, to a regulatory polynucleotide.

30. A method for modulating liver cell proliferation or tumorigenesis, comprising contacting a liver cell with an agent for a time and under conditions sufficient to modulate the level and/or functional activity of an *IGFBP* expression product, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of a gene selected from an *IGFBP* gene or a gene relating to the same regulatory or biosynthetic pathway as the *IGFBP* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a change in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a change in the level of expression of said marker gene.

31. A method for modulating liver cell proliferation or tumorigenesis, comprising contacting a liver cell with an agent for a time and under conditions sufficient to reduce the level and/or functional activity of an *IGFBP-1* expression product, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–1* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

32. A method for modulating liver cell proliferation or tumorigenesis, comprising contacting a liver cell with an agent for a time and under conditions sufficient to reduce the level and/or functional activity of an *IGFBP-2* expression product, wherein said

agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–2* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting a reduction in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or a reduction in the level and/or functional activity of an expression product of said marker gene.

33. A method for modulating liver cell proliferation or tumorigenesis, comprising contacting a liver cell with an agent for a time and under conditions sufficient to increase the level and/or functional activity of an *IGFBP-3* expression product, wherein said agent is optionally formulated with a pharmaceutically acceptable carrier and is identifiable by a screening assay comprising:

– contacting a preparation comprising at least a portion of an expression product of an *IGFBP–3* gene, or a variant or derivative of said expression product, or comprising a genetic sequence which regulates expression of said expression product and which is operably connected to a marker gene, with said agent; and

– detecting an increase in the level and/or functional activity of said at least a portion of said expression product or said variant or derivative or an increase in the level and/or functional activity of an expression product of said marker gene.

34. A method for treatment and/or prophylaxis of a liver cancer or related condition, comprising administering to a patient in need of such treatment an effective amount of an agent that modulates the level and/or functional activity of an *IGFBP* expression product.

35. A method for treatment and/or prophylaxis of a liver cancer or related condition, comprising administering to a patient in need of such treatment an effective amount of an agent that decreases the level and/or functional activity of an *IGFBP–1* expression product.

36. A method for treatment and/or prophylaxis of a liver cancer or related condition, comprising administering to a patient in need of such treatment an effective amount of

an agent that decreases the level and/or functional activity of an *IGFBP–2* expression product.

37. A method for treatment and/or prophylaxis of a liver cancer or related condition, comprising administering to a patient in need of such treatment an effective amount of an agent that increases the level and/or functional activity of an *IGFBP–3* expression product.

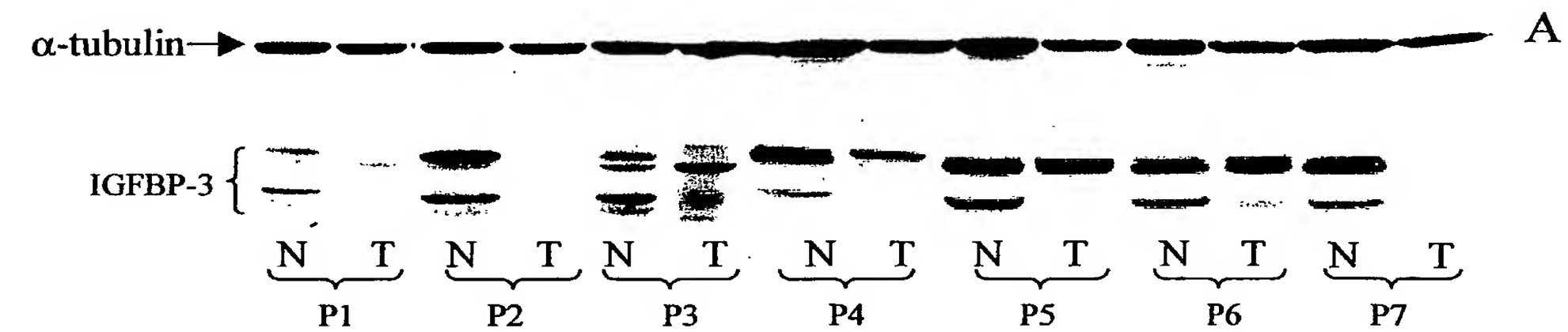
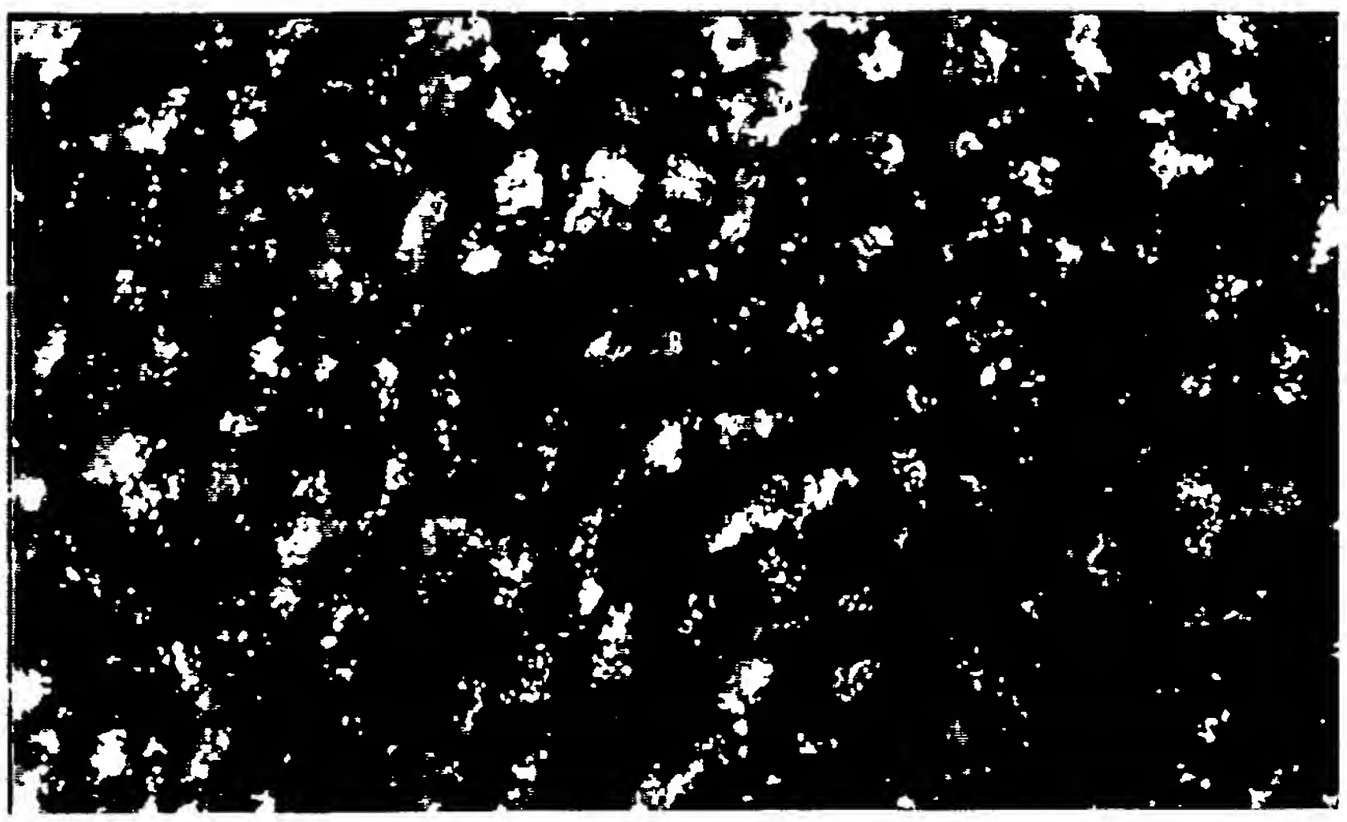

α-tubulin
A
IGFBP-3
N T N T N T N T N T N T N T
P1 P2 P3 P4 P5 P6 P7
α-tubulin
B
IGFBP-3
N T N T N T N T N T N T N T
P8 P9 P10 P11 P12 P13 P14
α-tubulin
C
IGFBP-3
IGFBP-3 Fragment
N T N T N T
Serum
P15 P16 P17

FIGURE 1

A



B

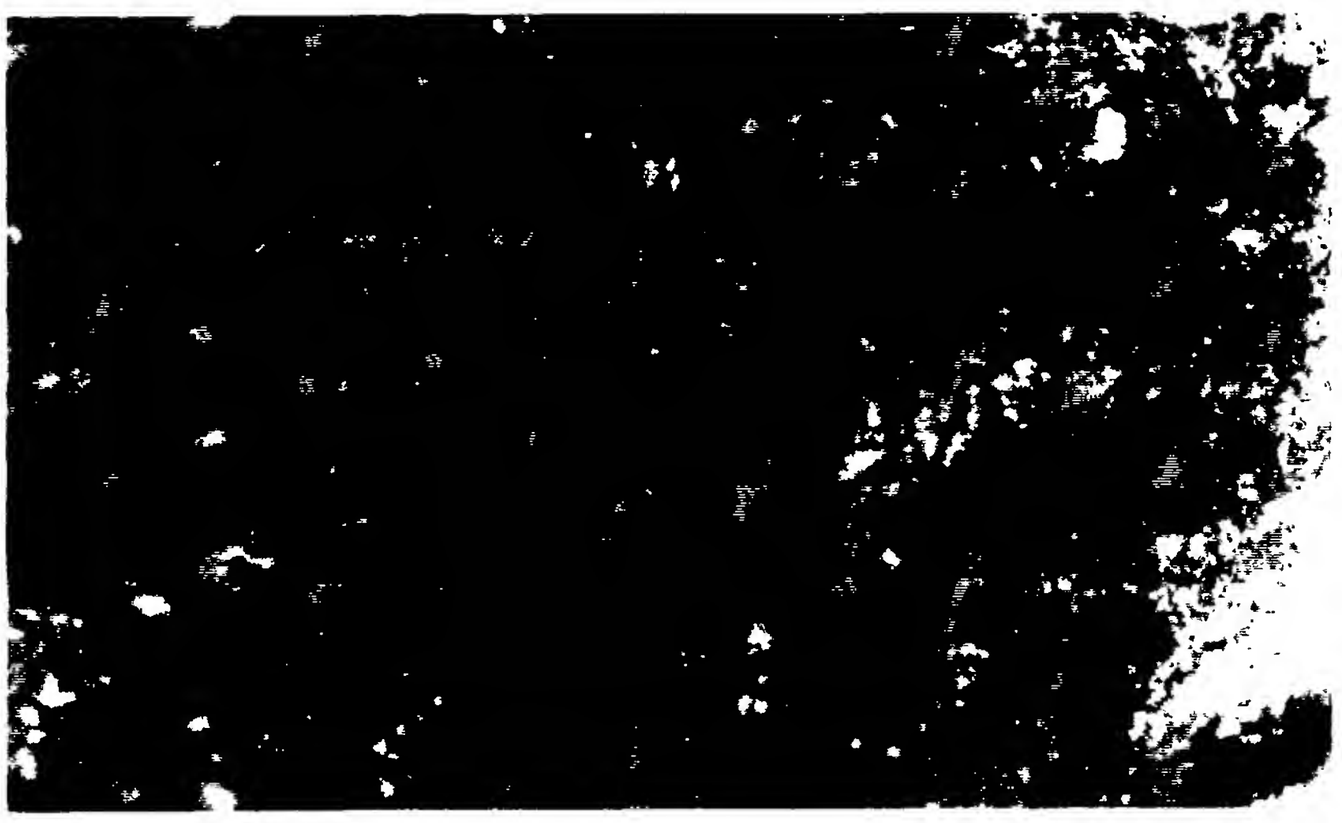
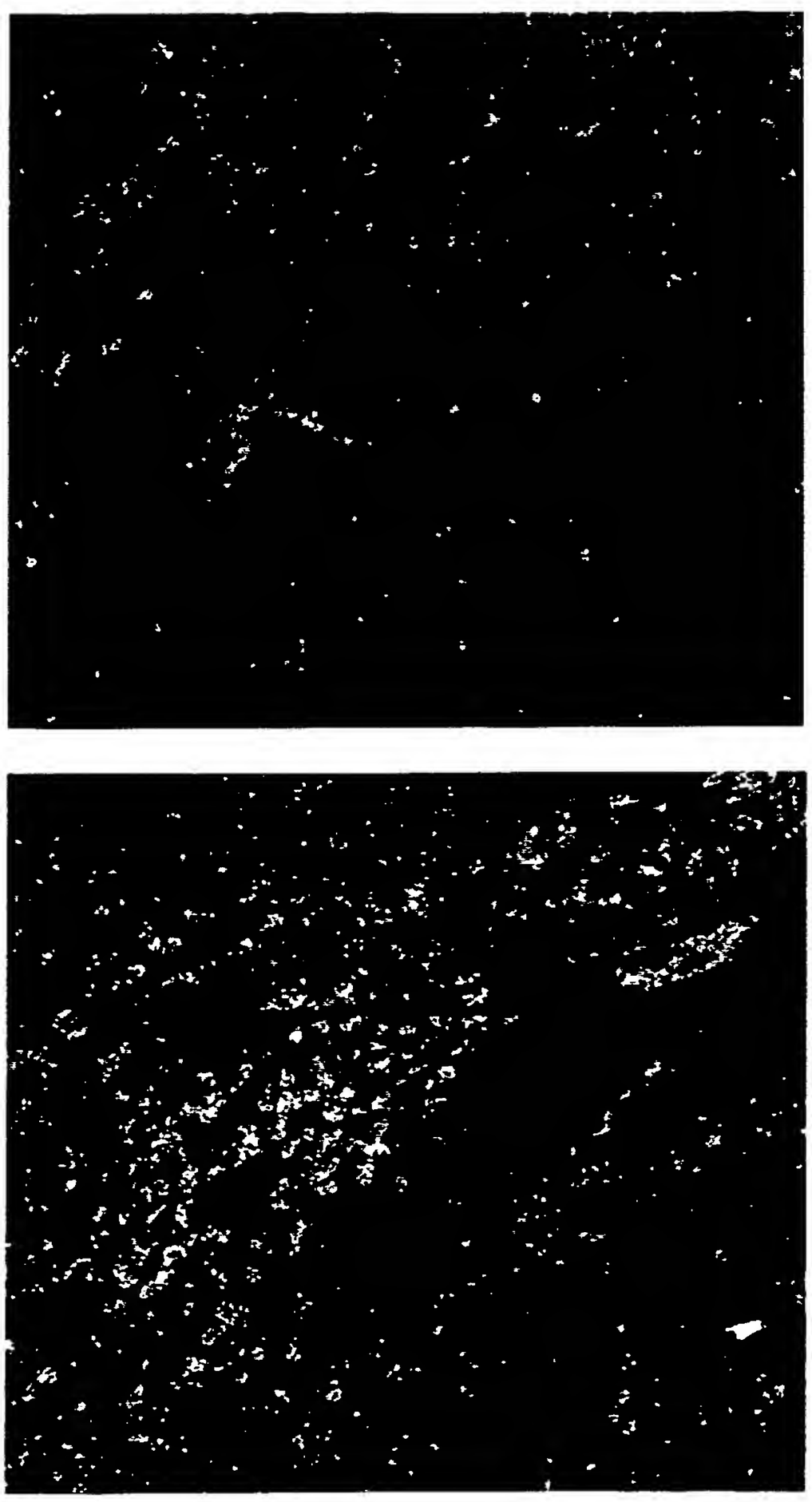


FIGURE 2

FIGURE 3



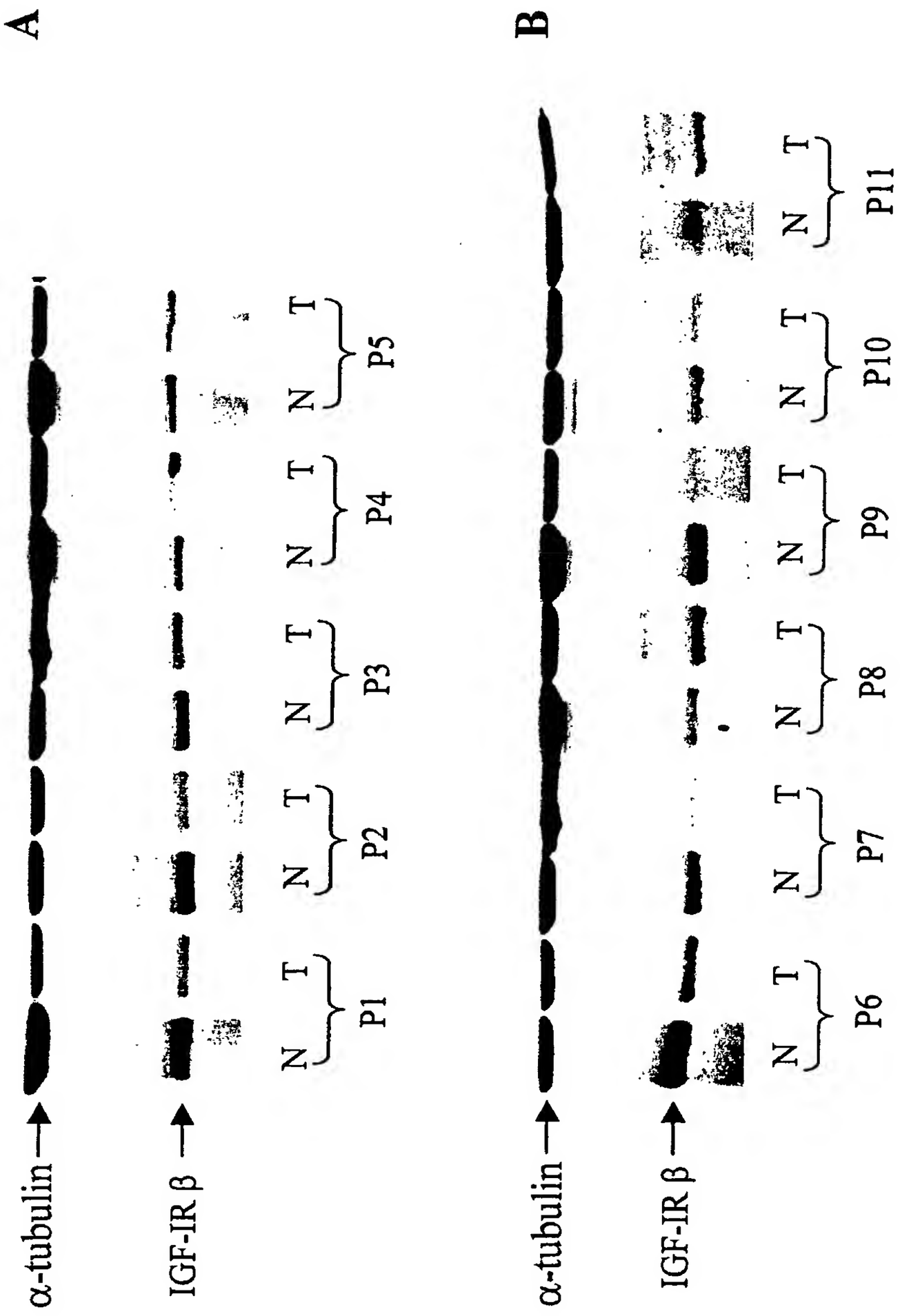

A
B
α-tubulin
IGF-IR β
N T
P1
N T
P2
N T
P3
N T
P4
N T
P5
N T
P6
N T
P7
N T
P8
N T
P9
N T
P10
N T
P11

FIGURE 4

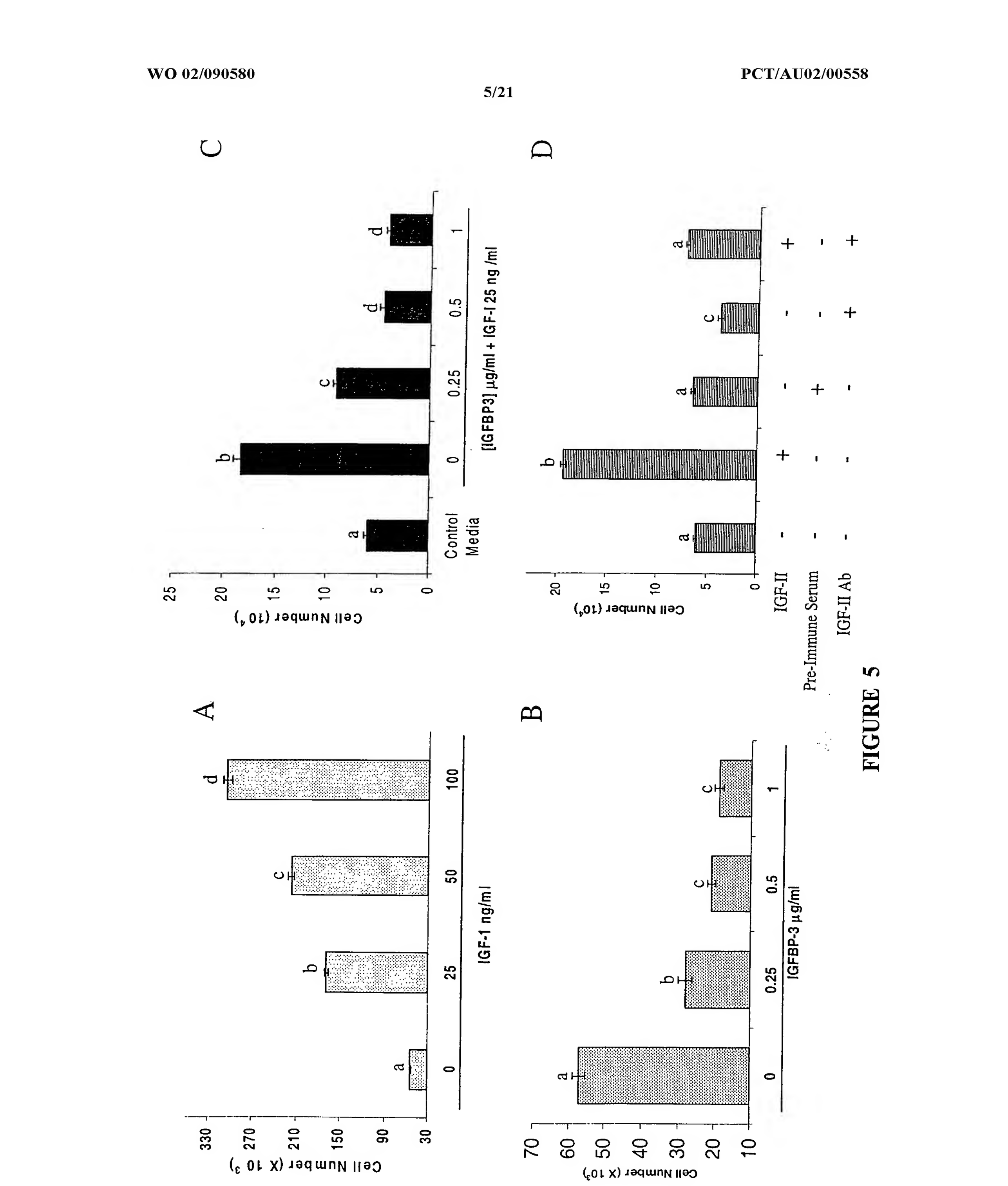

A
d
c
b
a
Cell Number (X 10 3)
330
270
210
150
90
30
0
25
50
100
IGF-1 ng/ml
B
a
c
b
c
Cell Number (X 10³)
70
60
50
40
30
20
10
0
0.25
0.5
1
IGFBP-3 µg/ml
C
d
d
c
b
a
Cell Number (10⁴)
25
20
15
10
5
0
0.25
0.5
1
[IGFBP3] µg/ml + IGF-I 25 ng /ml
Control Media
D
a
c
a
b
a
Cell Number (10⁴)
20
15
10
5
0
+ - +
- - +
- + -
+ - -
IGF-II -
Pre-Immune Serum -
IGF-II Ab -

FIGURE 5

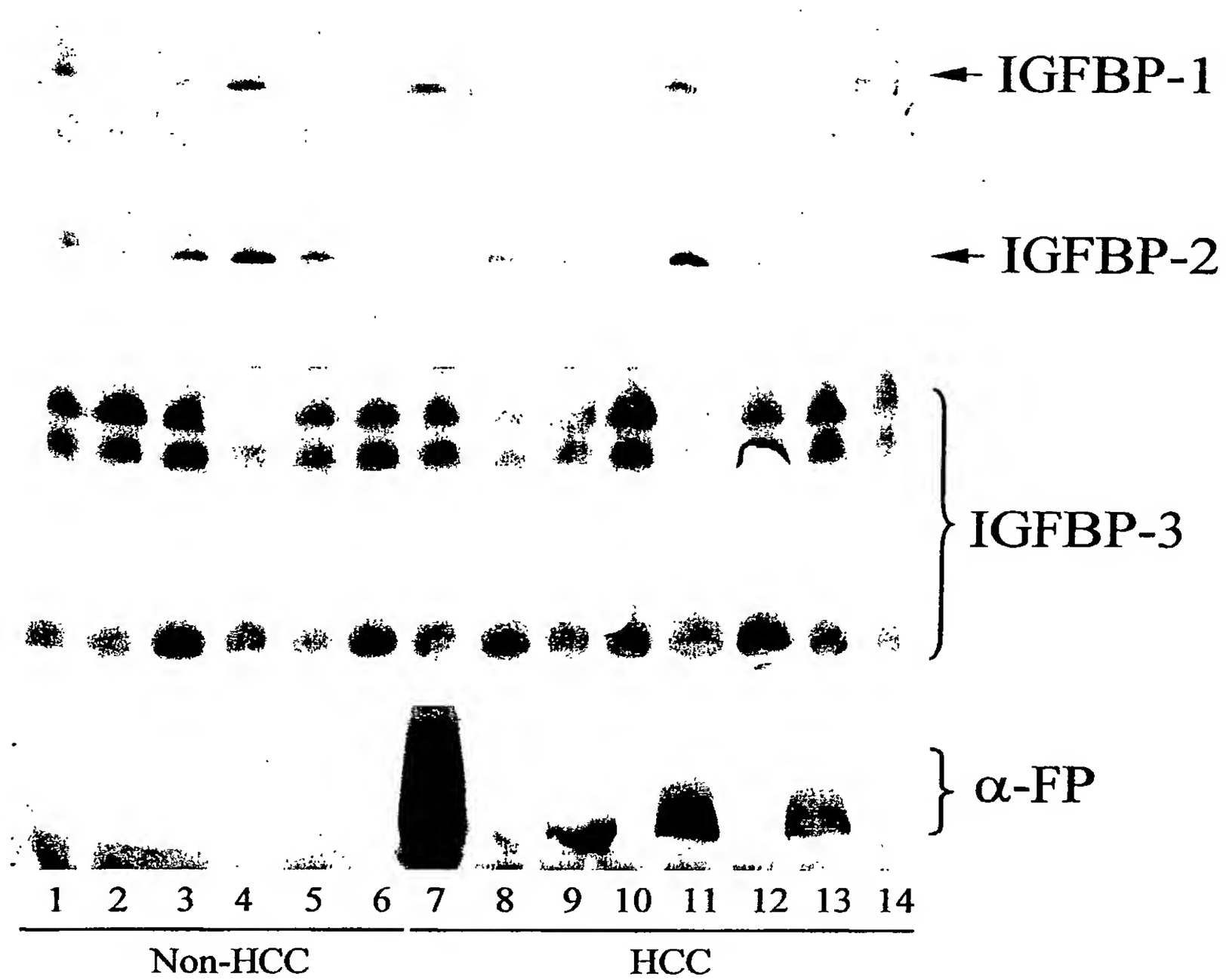

19/12/00 Serum1
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6A

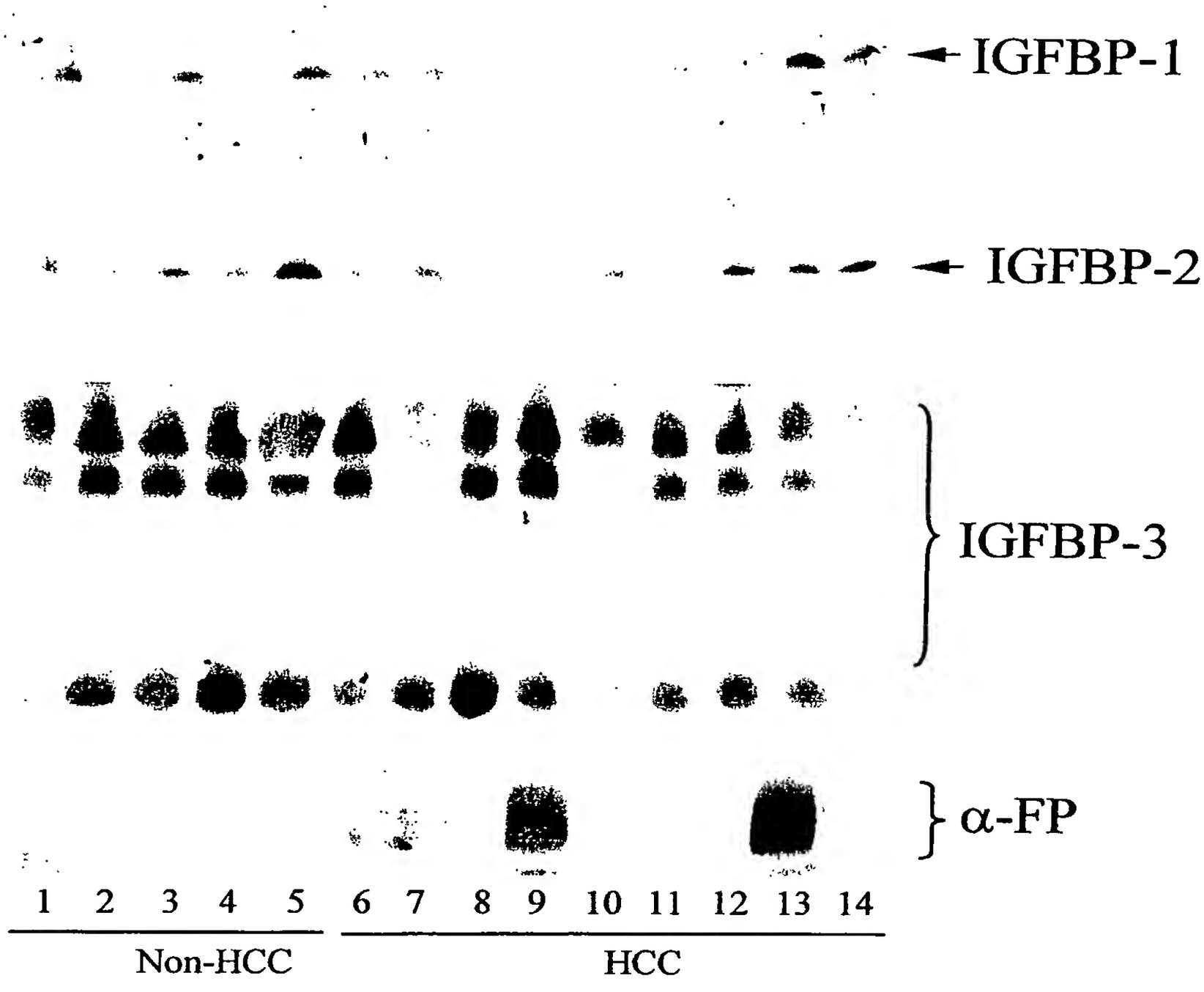

19/12/00 Serum 2
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6B

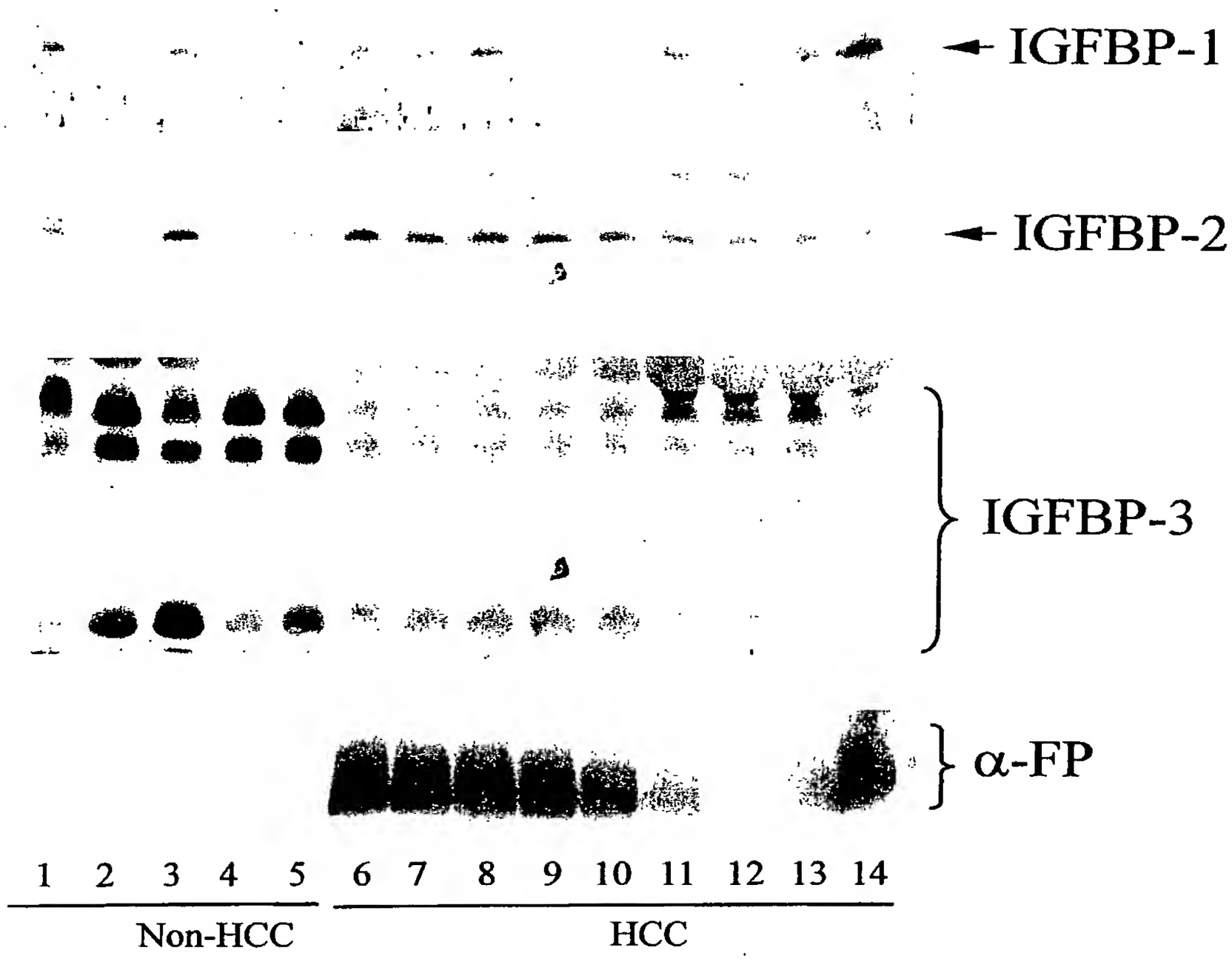

19/12/00 Serum 3
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6C

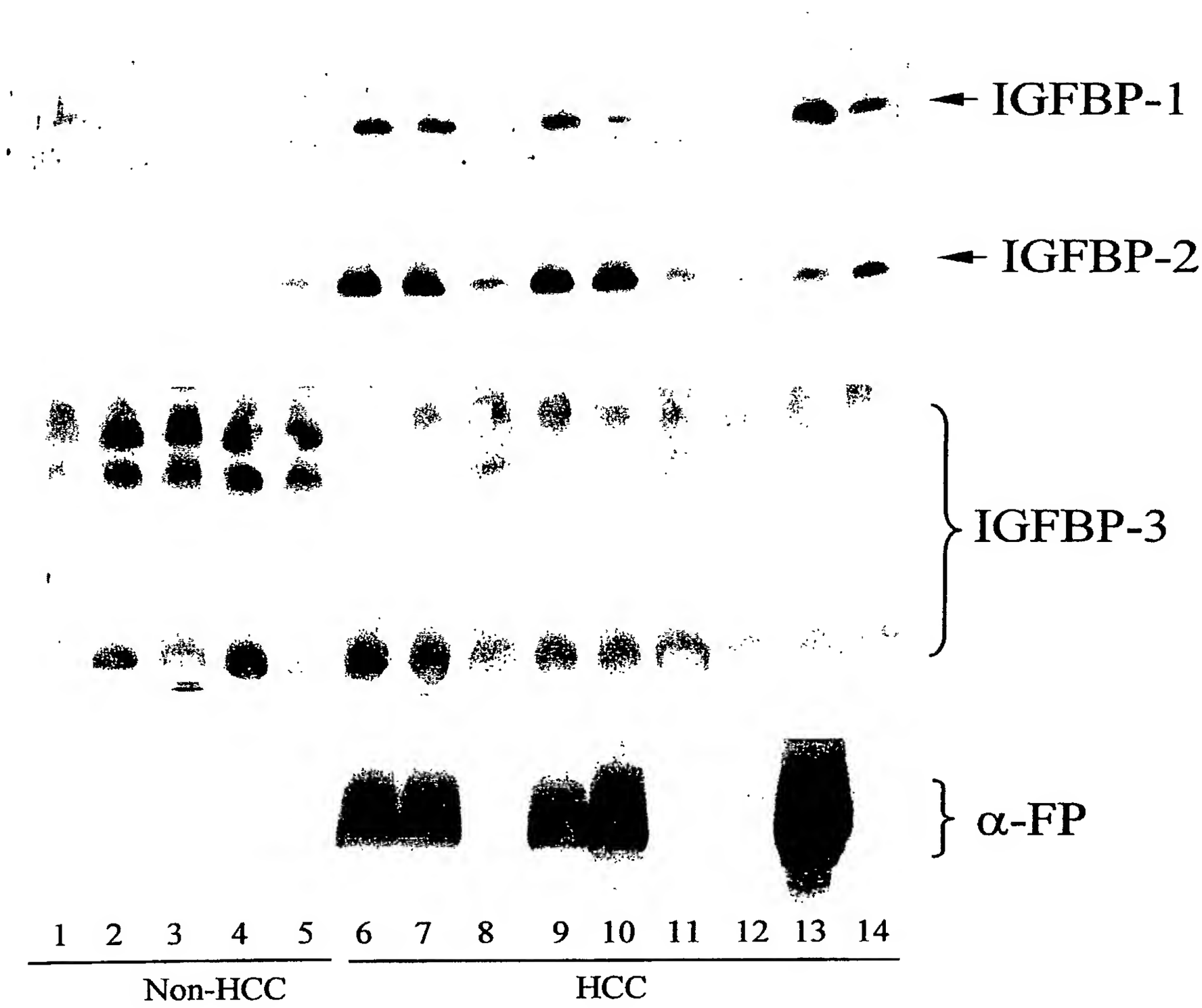

19/12/00 Serum 4
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6D

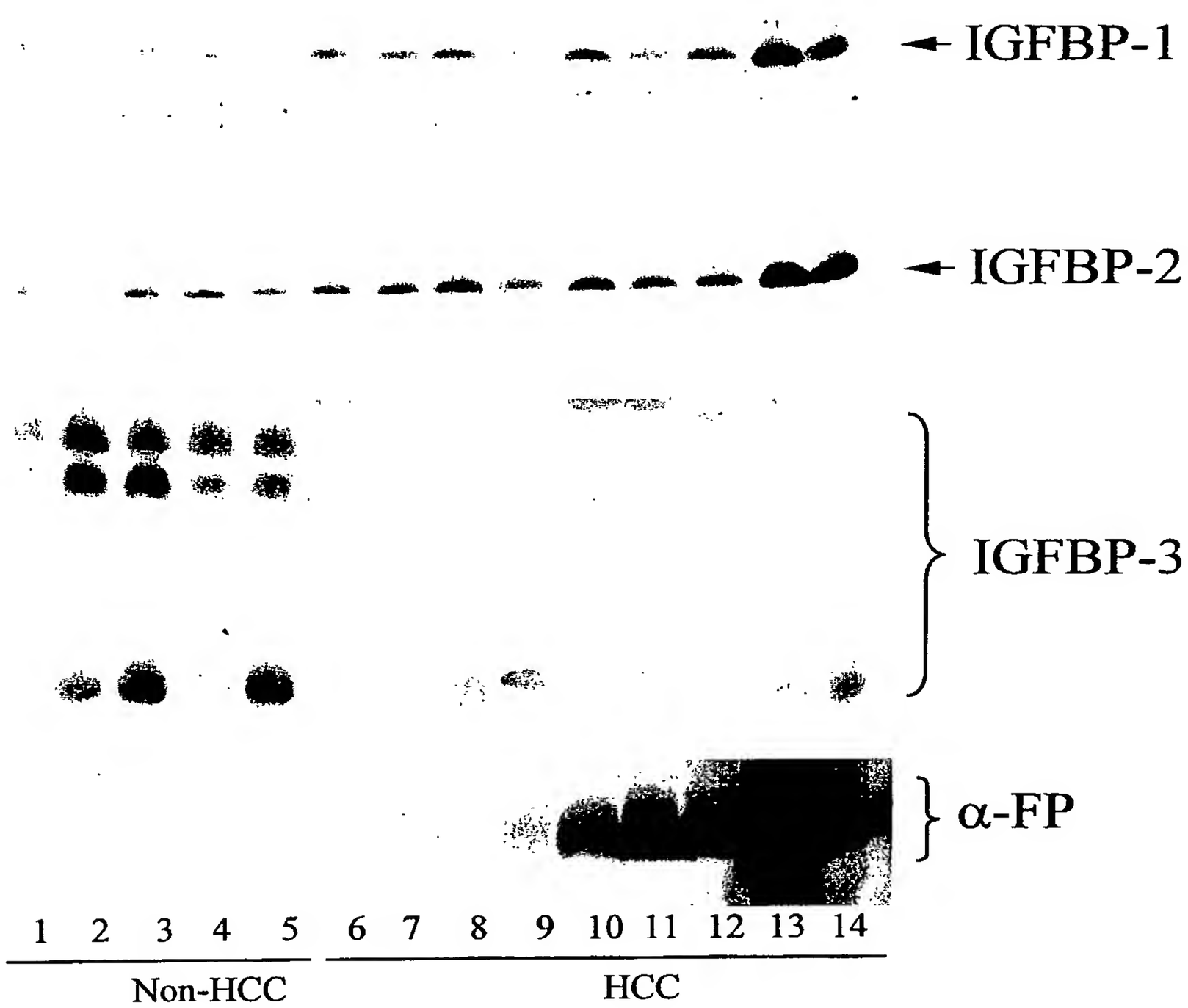

19/12/00 Serum 5
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6E

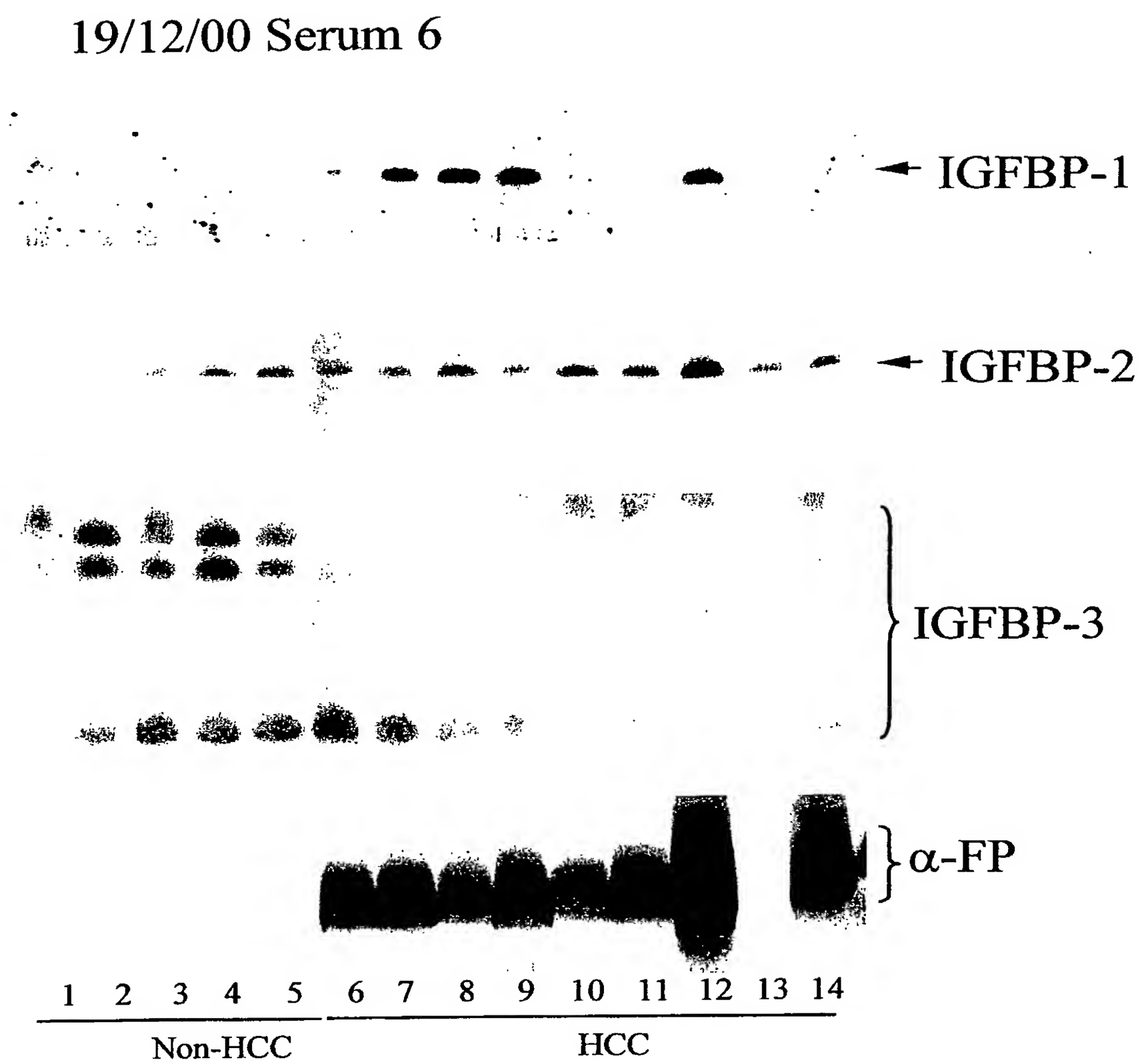

19/12/00 Serum 6
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6F

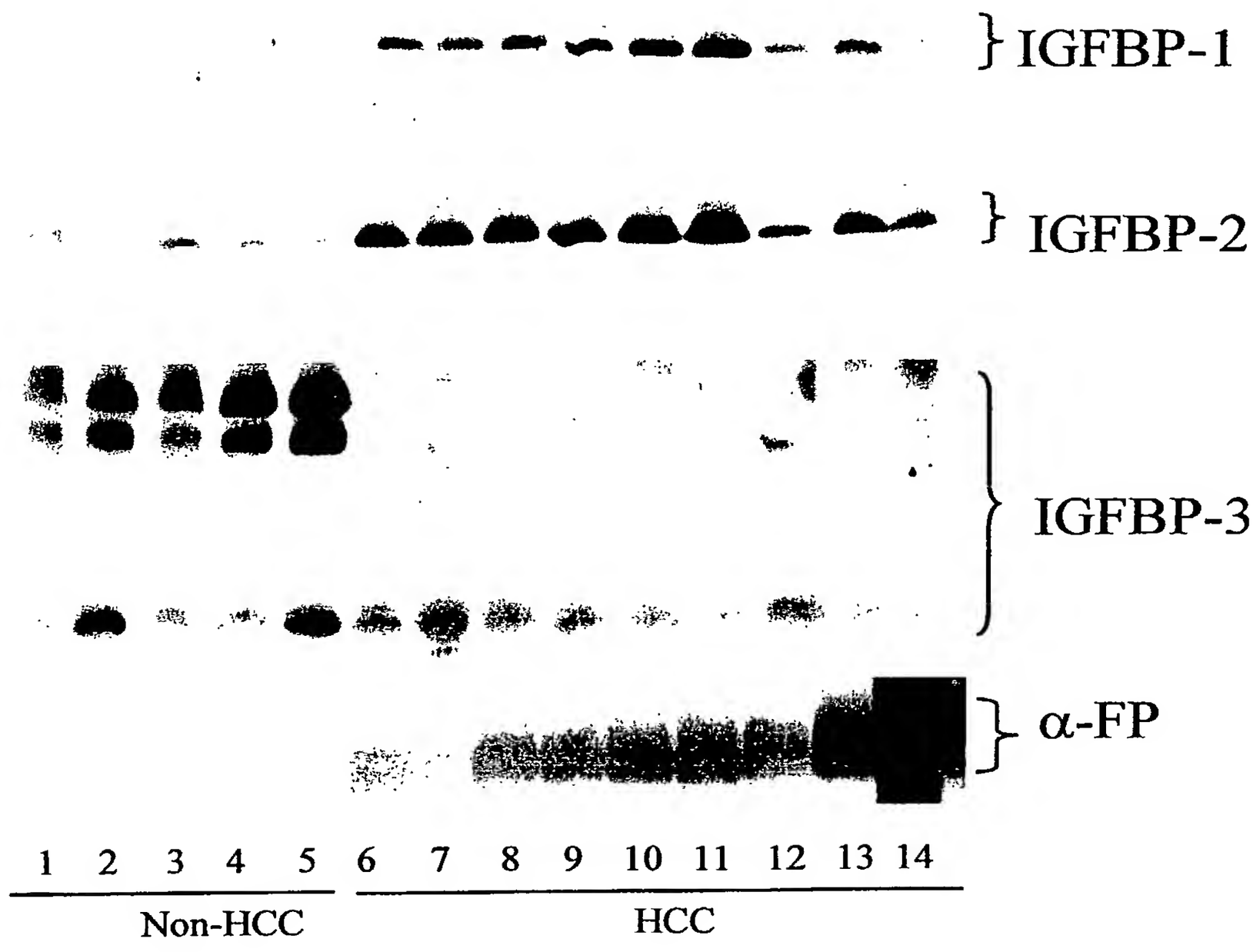

19/12/00 Serum7
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6G

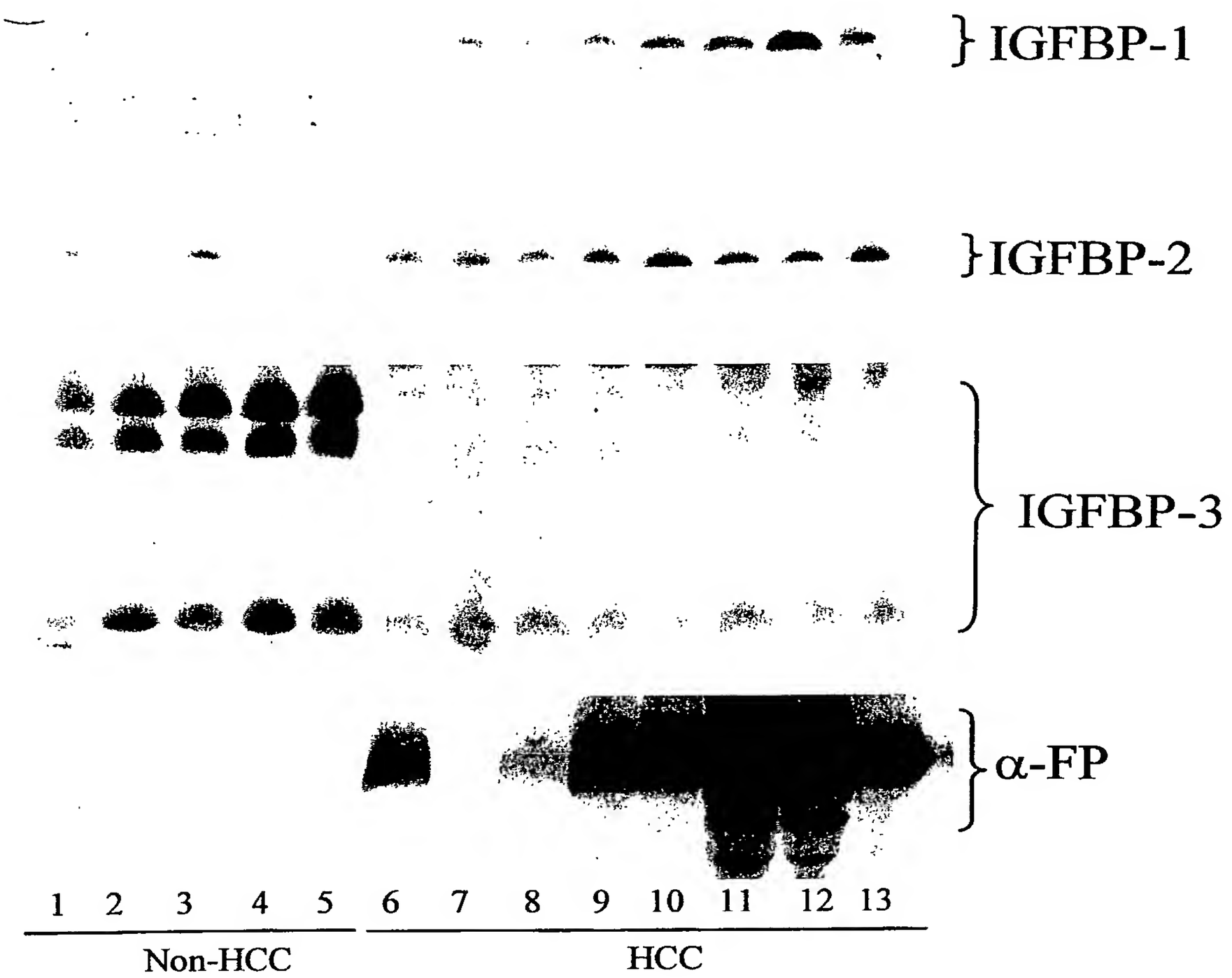

19/12/00 Serum 8
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13
Non-HCC
HCC

FIGURE 6H

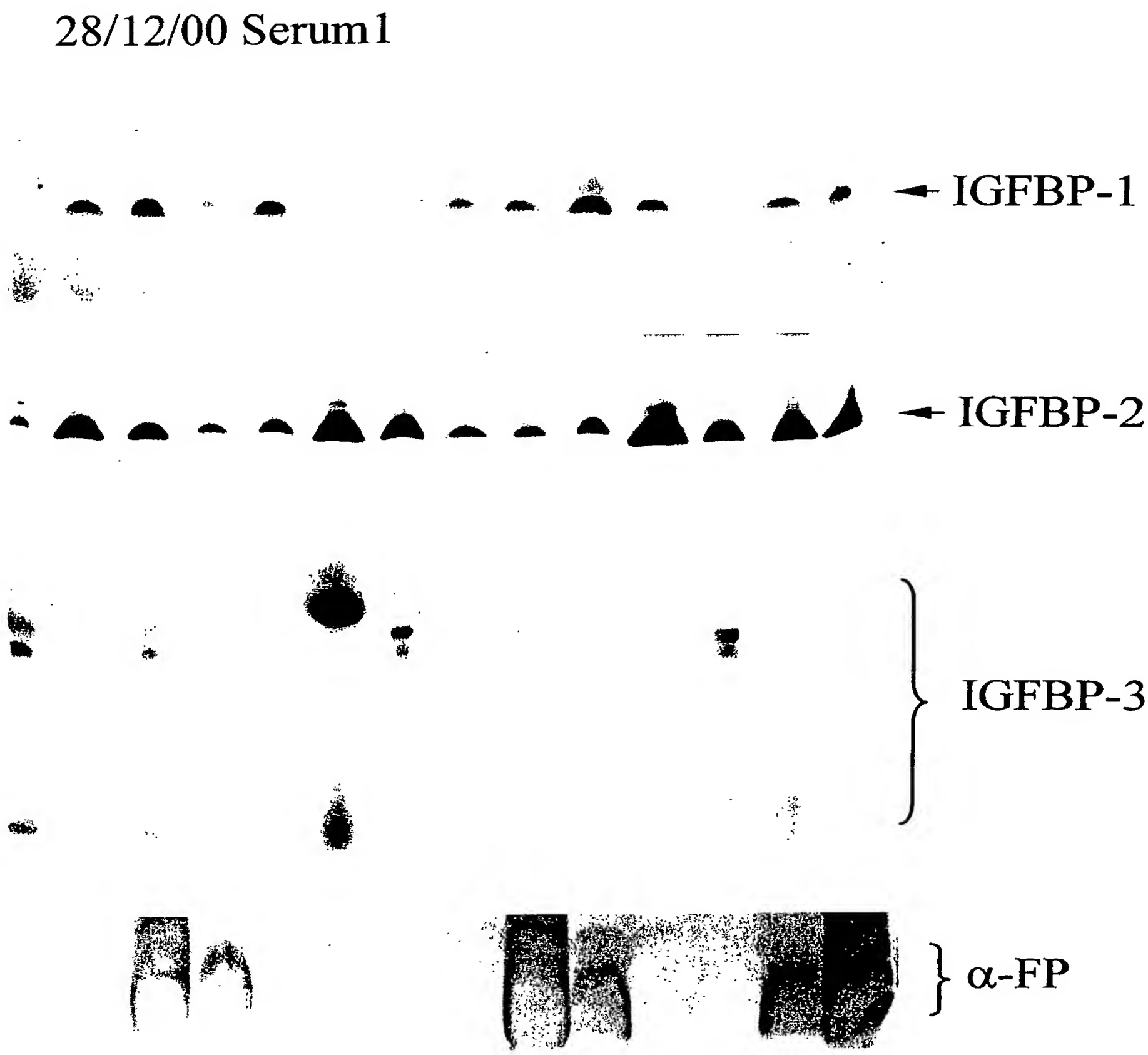

28/12/00 Serum1
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6I

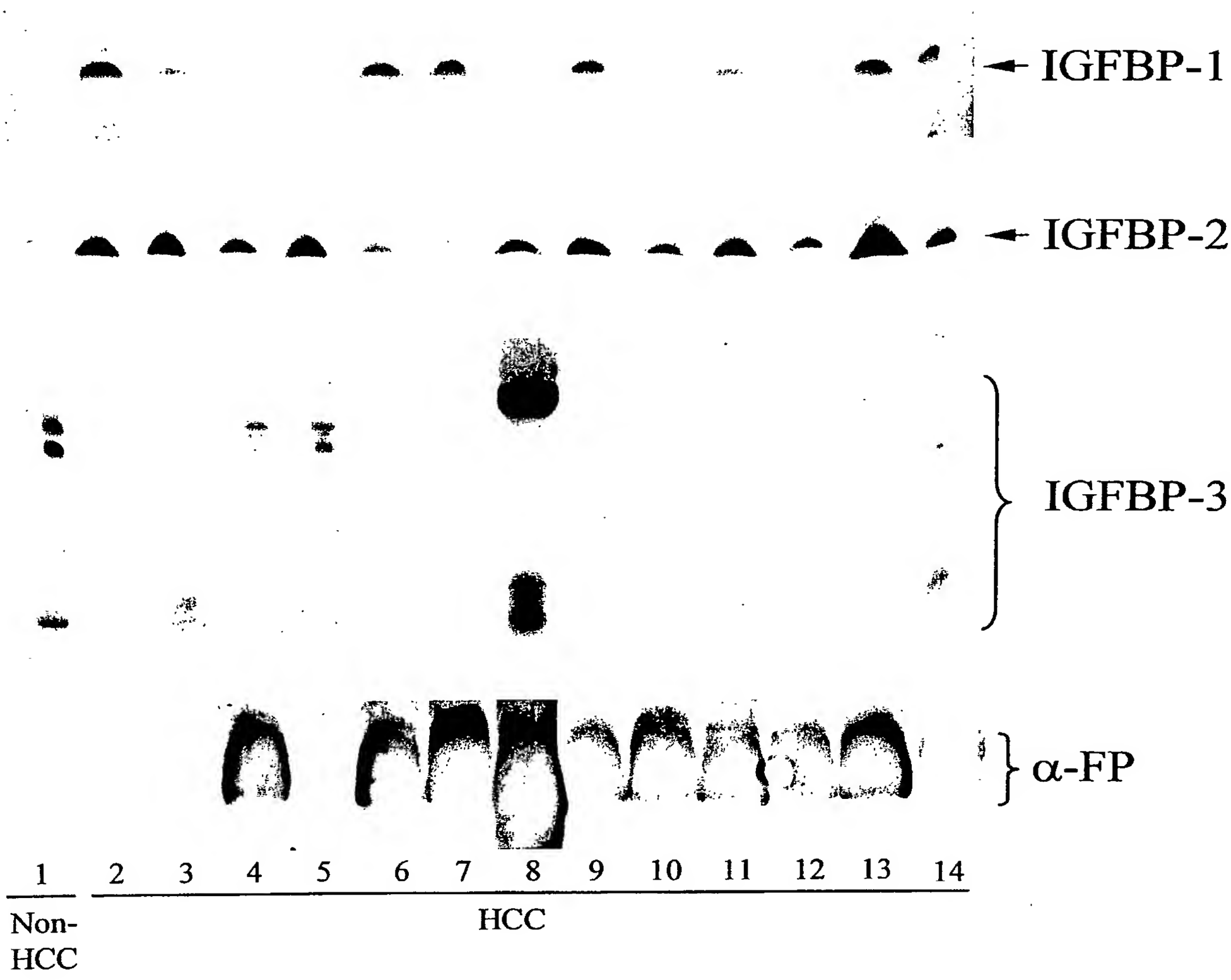

28/12/00 Serum 2
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1
2
3
4
5
6
7
8
9
10
11
12
13
14
Non-HCC
HCC

FIGURE 6J

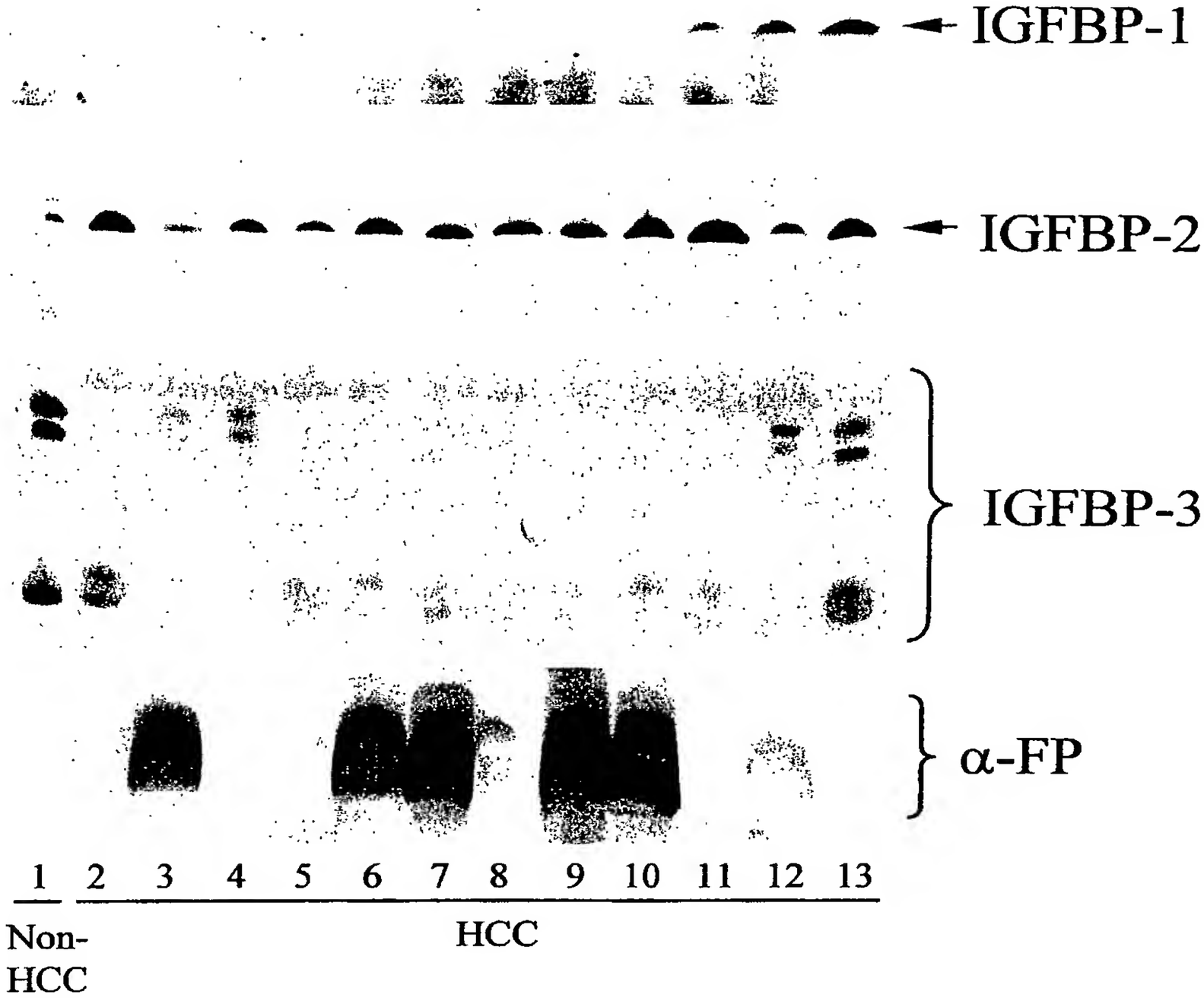

28/12/00 Serum 3
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13
Non-HCC
HCC

FIGURE 6K

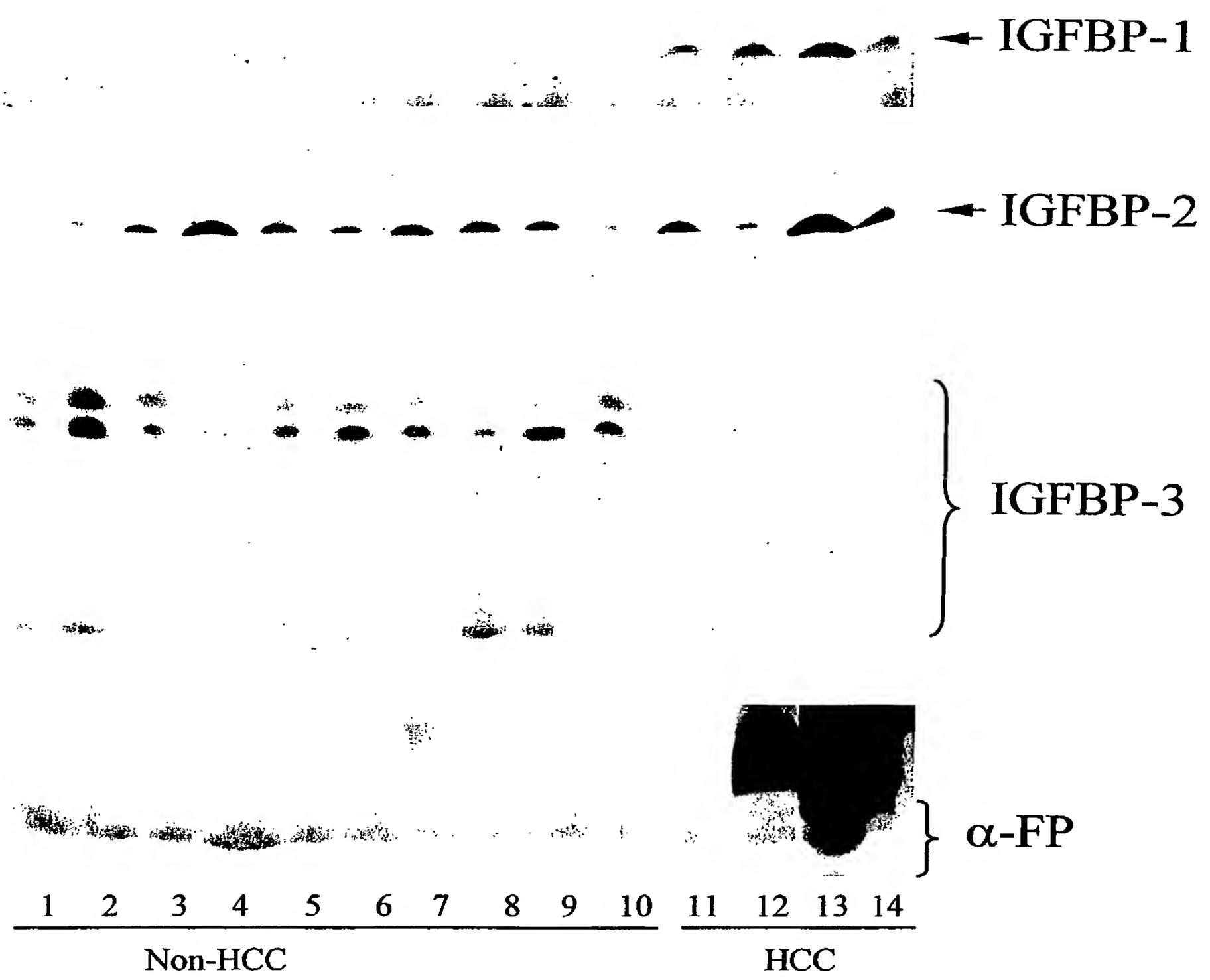
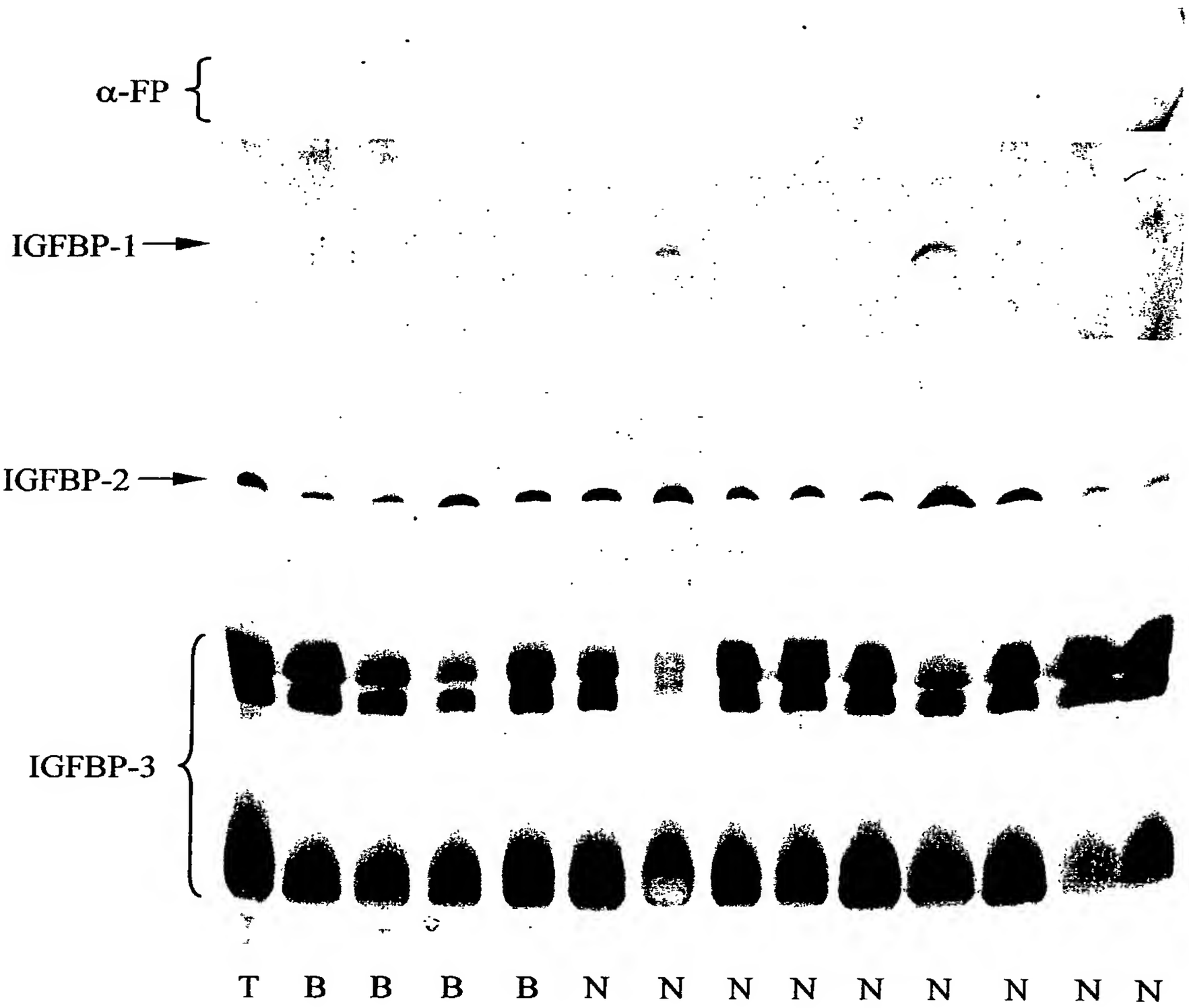

28/12/00 Serum 6
IGFBP-1
IGFBP-2
IGFBP-3
α-FP
1 2 3 4 5 6 7 8 9 10 11 12 13 14
Non-HCC
HCC

FIGURE 6L



N: Normal
T: Prostate Cancer
B: Benign Prostate Hyperplasia

FIGURE 6M

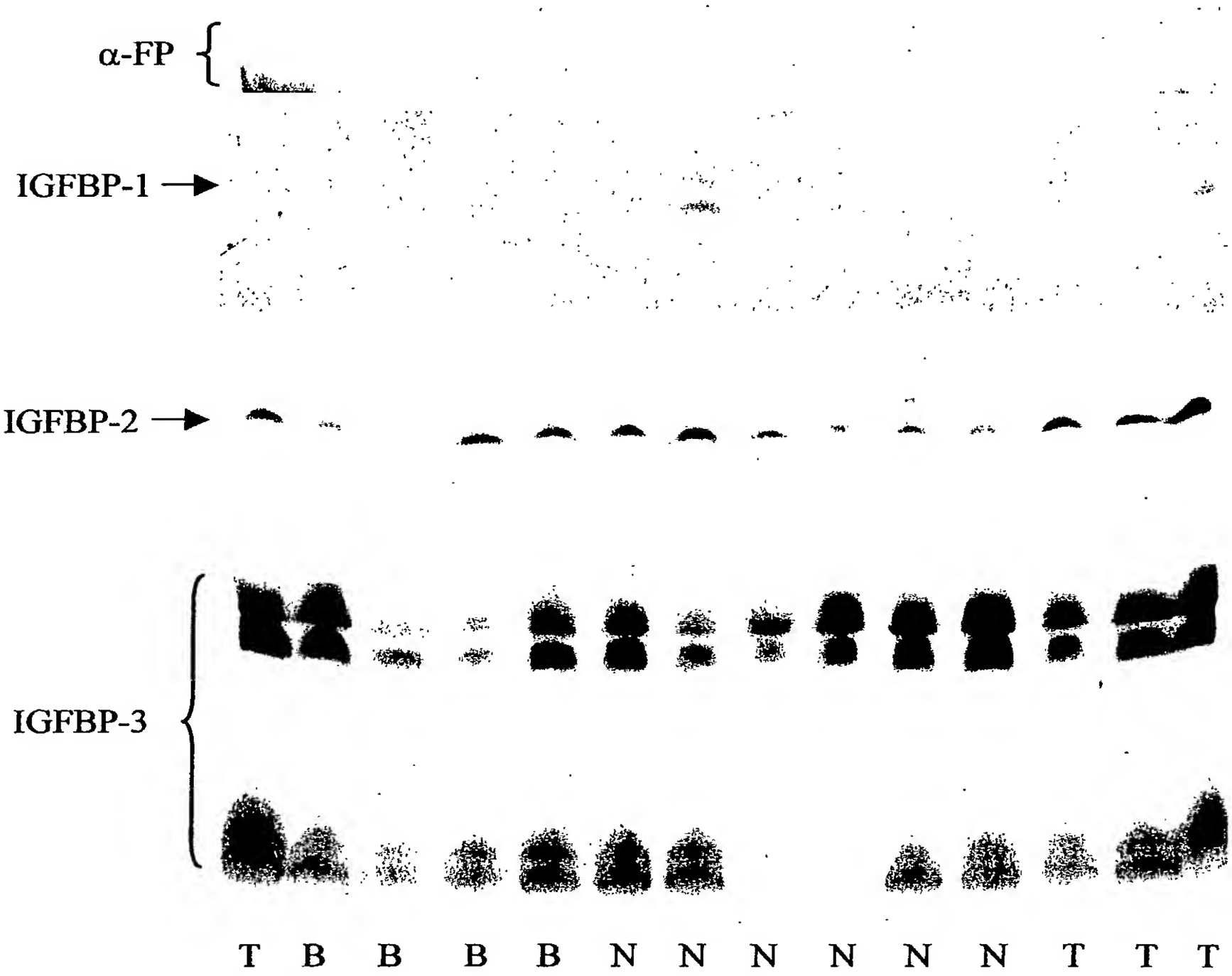
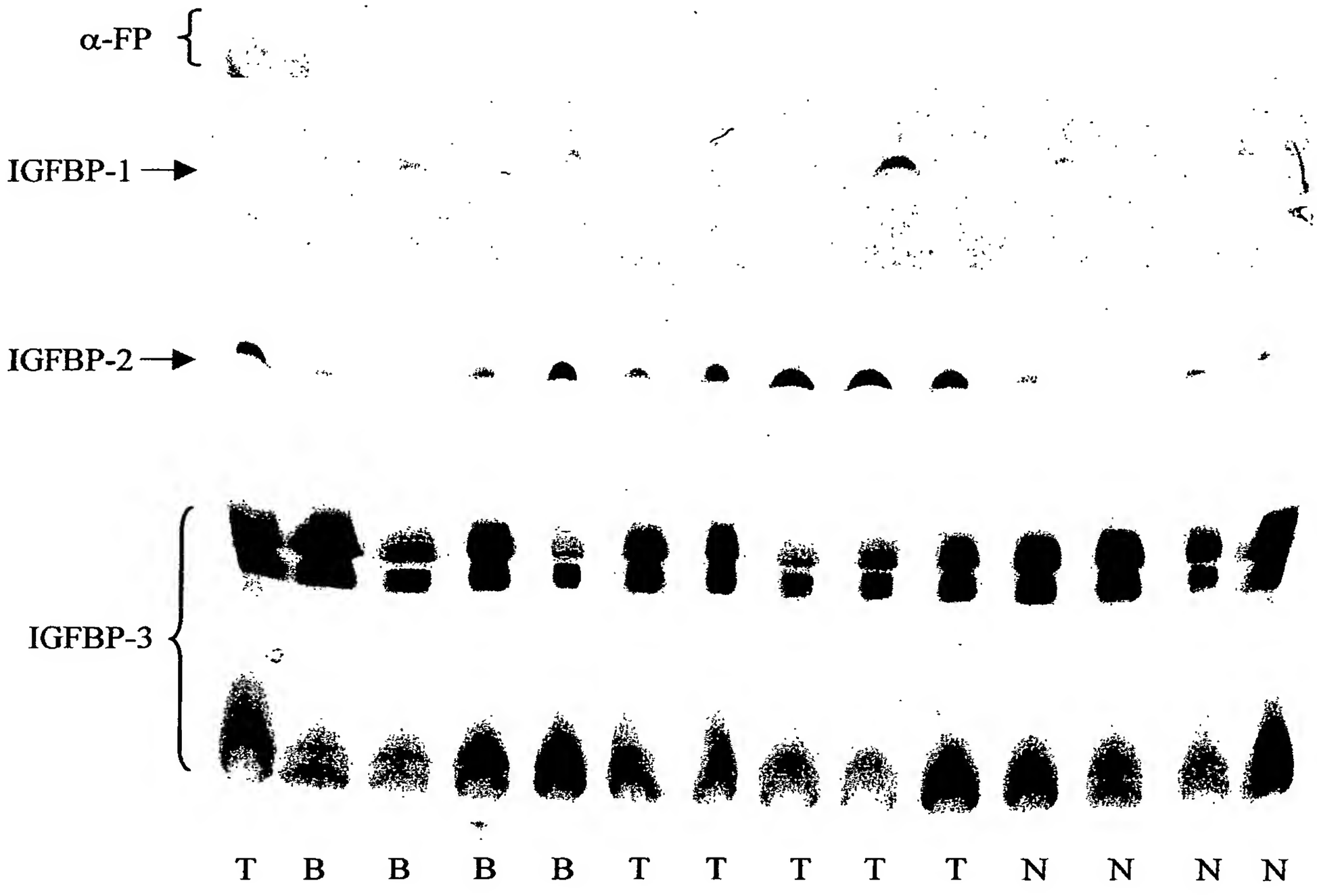
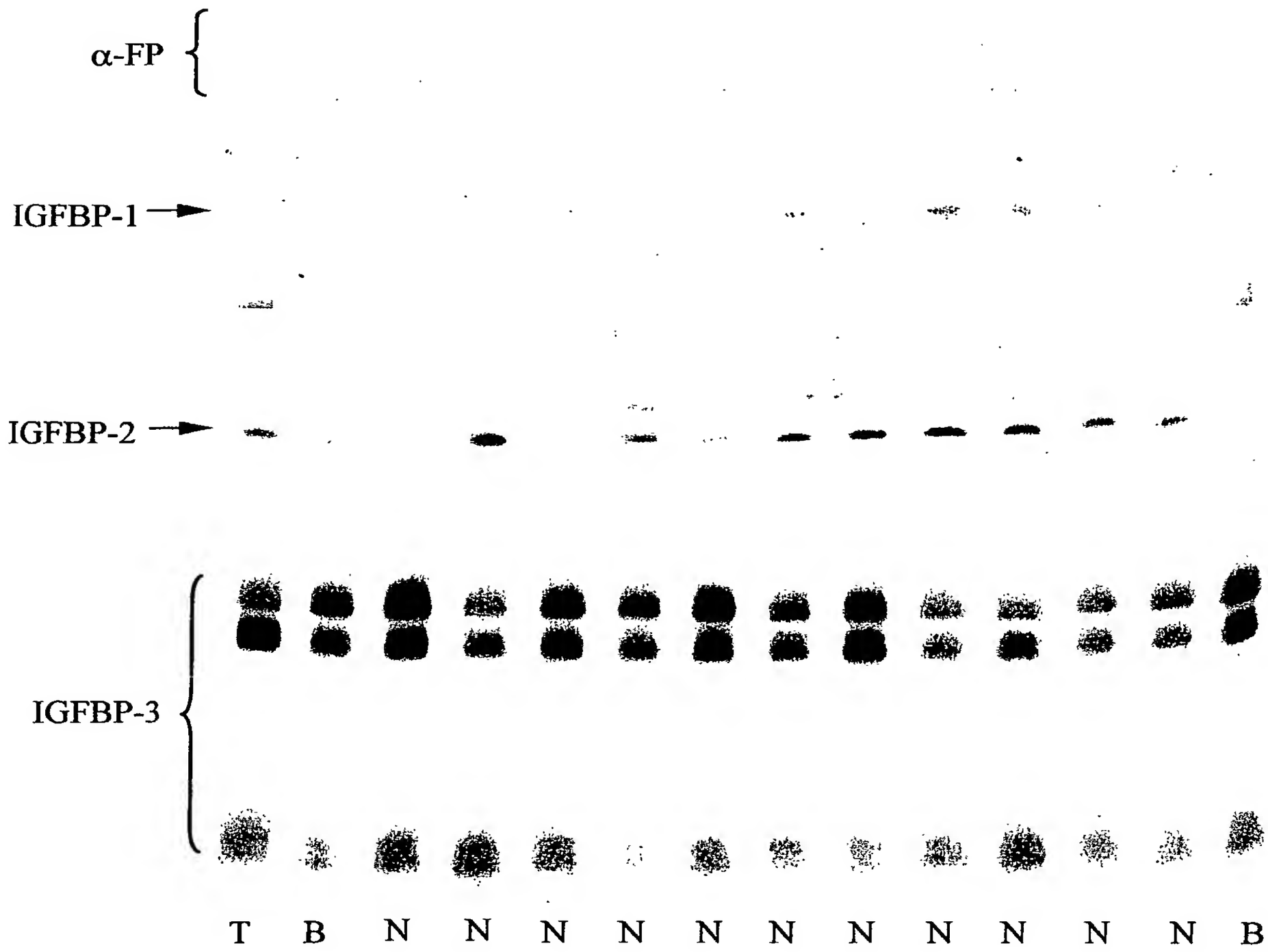

α-FP
IGFBP-1
IGFBP-2
IGFBP-3
T B B B B N N N N N N T T T
N: Normal
T: Prostate Cancer
B: Benign Prostate Hyperplasia

FIGURE 6N



N: Normal
T: Prostate Cancer
B: Benign Prostate Hyperplasia

FIGURE 6O



N: Normal
T: Prostate Cancer
B: Benign Prostate Hyperplasia

FIGURE 6P

SEQUENCE LISTING

<110> National Cancer Centre of Singapore (all states, except U.S.)
Singapore General Hospital Pte Ltd (all states, except U.S.)
Huynh, The Hung (U.S. only)
Chow, Pierce Kah Hoe (U.S. only)
Soo, Khee Chee (U.S. only)

<120> Polynucleotides and polypeptides linked to cancer

<130> 2521505/VPA

<140> Yet to be assigned
<141> 2002-05-03

<150> U.S. 60/288,441
<151> 2001-05-03

<160> 12

<170> PatentIn version 3.1

<210> 1
<211> 1514
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (166)..(945)
<223>

<400> 1
atcggccacc gccatcccat ccagcgagca tctgccgccg cgccgccgcc accctcccag 60

agagcactgg ccaccgctcc accatcactt gcccagagtt tgggccaccg cccgccgcca 120

ccagcccaga gagcatcggc ccctgtctgc tgctcgcgcc tggag atg tca gag gtc 177
Met Ser Glu Val
1

ccc gtt gct cgc gtc tgg ctg gta ctg ctc ctg ctg act gtc cag gtc 225
Pro Val Ala Arg Val Trp Leu Val Leu Leu Leu Leu Thr Val Gln Val
5 10 15 20

ggc gtg aca gcc ggc gct ccg tgg cag tgc gcg ccc tgc tcc gcc gag 273
Gly Val Thr Ala Gly Ala Pro Trp Gln Cys Ala Pro Cys Ser Ala Glu
25 30 35

aag ctc gcg ctc tgc ccg ccg gtg tcc gcc tcg tgc tcg gag gtc acc 321
Lys Leu Ala Leu Cys Pro Pro Val Ser Ala Ser Cys Ser Glu Val Thr
40 45 50

cgg tcc gcc ggc tgc ggc tgt tgc ccg atg tgc gcc ctg cct ctg ggc 369
Arg Ser Ala Gly Cys Gly Cys Cys Pro Met Cys Ala Leu Pro Leu Gly
55 60 65

gcc gcg tgc ggc gtg gcg act gca cgc tgc gcc cgg gga ctc agt tgc 417
Ala Ala Cys Gly Val Ala Thr Ala Arg Cys Ala Arg Gly Leu Ser Cys
70 75 80

```
cgc gcg ctg ccg ggg gag cag caa cct ctg cac gcc ctc acc cgc ggc      465
Arg Ala Leu Pro Gly Glu Gln Gln Pro Leu His Ala Leu Thr Arg Gly
85                  90                  95                  100

caa ggc gcc tgc gtg cag gag tct gac gcc tcc gct ccc cat gct gca      513
Gln Gly Ala Cys Val Gln Glu Ser Asp Ala Ser Ala Pro His Ala Ala
                105                 110                 115

gag gca ggg agc cct gaa agc cca gag agc acg gag ata act gag gag      561
Glu Ala Gly Ser Pro Glu Ser Pro Glu Ser Thr Glu Ile Thr Glu Glu
            120                 125                 130

gag ctc ctg gat aat ttc cat ctg atg gcc cct tct gaa gag gat cat      609
Glu Leu Leu Asp Asn Phe His Leu Met Ala Pro Ser Glu Glu Asp His
        135                 140                 145

tcc atc ctt tgg gac gcc atc agt acc tat gat ggc tcg aag gct ctc      657
Ser Ile Leu Trp Asp Ala Ile Ser Thr Tyr Asp Gly Ser Lys Ala Leu
    150                 155                 160

cat gtc acc aac atc aaa aaa tgg aag gag ccc tgc cga ata gaa ctc      705
His Val Thr Asn Ile Lys Lys Trp Lys Glu Pro Cys Arg Ile Glu Leu
165                 170                 175                 180

tac aga gtc gta gag agt tta gcc aag gca cag gag aca tca gga gaa      753
Tyr Arg Val Val Glu Ser Leu Ala Lys Ala Gln Glu Thr Ser Gly Glu
                185                 190                 195

gaa att tcc aaa ttt tac ctg cca aac tgc aac aag aat gga ttt tat      801
Glu Ile Ser Lys Phe Tyr Leu Pro Asn Cys Asn Lys Asn Gly Phe Tyr
            200                 205                 210

cac agc aga cag tgt gag aca tcc atg gat gga gag gcg gga ctc tgc      849
His Ser Arg Gln Cys Glu Thr Ser Met Asp Gly Glu Ala Gly Leu Cys
        215                 220                 225

tgg tgc gtc tac cct tgg aat ggg aag agg atc cct ggg tct cca gag      897
Trp Cys Val Tyr Pro Trp Asn Gly Lys Arg Ile Pro Gly Ser Pro Glu
    230                 235                 240

atc agg gga gac ccc aac tgc cag ata tat ttt aat gta caa aac tga      945
Ile Arg Gly Asp Pro Asn Cys Gln Ile Tyr Phe Asn Val Gln Asn
245                 250                 255

aaccagatga aataatgttc tgtcacgtga aatatttaag tatatagtat atttatactc   1005

tagaacatgc acatttatat atatatgtat atgtatatat atatagtaac tacttttttat   1065

actccataca taacttgata tagaaagctg tttatttatt cactgtaagt ttatttttttc   1125

tacacagtaa aaacttgtac tatgttaata acttgtccta tgtcaatttg tatatcatga   1185

aacacttctc atcatattgt atgtaagtaa ttgcatttct gctcttccaa agctcctgcg   1245

tctgttttta aagagcatgg aaaaatactg cctagaaaat gcaaaatgaa ataagagaga   1305

gtagtttttc agctagtttg aaggaggacg gttaacttgt atattccacc attcacattt   1365

gatgtacatg tgtagggaaa gttaaaagtg ttgattacat aatcaaagct acctgtggtg   1425

atgttgccac ctgttaaaat gtacactgga tatgttgtta aacacgtgtc tataatggaa   1485
```

acatttacaa taaatattct gcatggaaa 1514

<210> 2
<211> 259
<212> PRT
<213> Homo sapiens

<400> 2

Met Ser Glu Val Pro Val Ala Arg Val Trp Leu Val Leu Leu Leu Leu
1 5 10 15

Thr Val Gln Val Gly Val Thr Ala Gly Ala Pro Trp Gln Cys Ala Pro
20 25 30

Cys Ser Ala Glu Lys Leu Ala Leu Cys Pro Pro Val Ser Ala Ser Cys
35 40 45

Ser Glu Val Thr Arg Ser Ala Gly Cys Gly Cys Cys Pro Met Cys Ala
50 55 60

Leu Pro Leu Gly Ala Ala Cys Gly Val Ala Thr Ala Arg Cys Ala Arg
65 70 75 80

Gly Leu Ser Cys Arg Ala Leu Pro Gly Glu Gln Gln Pro Leu His Ala
85 90 95

Leu Thr Arg Gly Gln Gly Ala Cys Val Gln Glu Ser Asp Ala Ser Ala
100 105 110

Pro His Ala Ala Glu Ala Gly Ser Pro Glu Ser Pro Glu Ser Thr Glu
115 120 125

Ile Thr Glu Glu Glu Leu Leu Asp Asn Phe His Leu Met Ala Pro Ser
130 135 140

Glu Glu Asp His Ser Ile Leu Trp Asp Ala Ile Ser Thr Tyr Asp Gly
145 150 155 160

Ser Lys Ala Leu His Val Thr Asn Ile Lys Lys Trp Lys Glu Pro Cys
165 170 175

Arg Ile Glu Leu Tyr Arg Val Val Glu Ser Leu Ala Lys Ala Gln Glu
180 185 190

Thr Ser Gly Glu Glu Ile Ser Lys Phe Tyr Leu Pro Asn Cys Asn Lys
195 200 205

```
Asn Gly Phe Tyr His Ser Arg Gln Cys Glu Thr Ser Met Asp Gly Glu
    210                 215                 220

Ala Gly Leu Cys Trp Cys Val Tyr Pro Trp Asn Gly Lys Arg Ile Pro
225                 230                 235                 240

Gly Ser Pro Glu Ile Arg Gly Asp Pro Asn Cys Gln Ile Tyr Phe Asn
                245                 250                 255

Val Gln Asn
```

<210> 3
<211> 1433
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (118)..(1104)
<223>

<400> 3

```
attcggggcg agggaggagg aagaagcgga ggaggcggct cccgctcgca gggccgtgca      60

cctgcccgcc cgcccgctcg ctcgctcgcc cgccgcgccg cgctgccgac cgccagc        117

atg ctg ccg aga gtg ggc tgc ccc gcg ctg ccg ctg ccg ccg ccg ccg       165
Met Leu Pro Arg Val Gly Cys Pro Ala Leu Pro Leu Pro Pro Pro Pro
1               5                   10                  15

ctg ctg ccg ctg ctg ccg ctg ctg ctg ctg cta ctg ggc gcg agt ggc       213
Leu Leu Pro Leu Leu Pro Leu Leu Leu Leu Leu Leu Gly Ala Ser Gly
            20                  25                  30

ggc ggc ggc ggg gcg cgc gcg gag gtg ctg ttc cgc tgc ccg ccc tgc       261
Gly Gly Gly Gly Ala Arg Ala Glu Val Leu Phe Arg Cys Pro Pro Cys
        35                  40                  45

aca ccc gag cgc ctg gcc gcc tgc ggg ccc ccg ccg gtt gcg ccg ccc       309
Thr Pro Glu Arg Leu Ala Ala Cys Gly Pro Pro Pro Val Ala Pro Pro
    50                  55                  60

gcc gcg gtg gcc gca gtg gcc gga ggc gcc cgc atg cca tgc gcg gag       357
Ala Ala Val Ala Ala Val Ala Gly Gly Ala Arg Met Pro Cys Ala Glu
65                  70                  75                  80

ctc gtc cgg gag ccg ggc tgc ggc tgc tgc tcg gtg tgc gcc cgg ctg       405
Leu Val Arg Glu Pro Gly Cys Gly Cys Cys Ser Val Cys Ala Arg Leu
                85                  90                  95

gag ggc gag gcg tgc ggc gtc tac acc ccg cgc tgc ggc cag ggg ctg       453
Glu Gly Glu Ala Cys Gly Val Tyr Thr Pro Arg Cys Gly Gln Gly Leu
            100                 105                 110

cgc tgc tat ccc cac ccg ggc tcc gag ctg ccc ctg cag gcg ctg gtc       501
```

```
Arg Cys Tyr Pro His Pro Gly Ser Glu Leu Pro Leu Gln Ala Leu Val
        115                 120                 125

atg ggc gag ggc act tgt gag aag cgc cgg gac gcc gag tat ggc gcc      549
Met Gly Glu Gly Thr Cys Glu Lys Arg Arg Asp Ala Glu Tyr Gly Ala
    130                 135                 140

agc ccg gag cag gtt gca gac aat ggc gat gac cac tca gaa gga ggc      597
Ser Pro Glu Gln Val Ala Asp Asn Gly Asp Asp His Ser Glu Gly Gly
145                 150                 155                 160

ctg gtg gag aac cac gtg gac agc acc atg aac atg ttg ggc ggg gga      645
Leu Val Glu Asn His Val Asp Ser Thr Met Asn Met Leu Gly Gly Gly
                165                 170                 175

ggc agt gct ggc cgg aag ccc ctc aag tcg ggt atg aag gag ctg gcc      693
Gly Ser Ala Gly Arg Lys Pro Leu Lys Ser Gly Met Lys Glu Leu Ala
            180                 185                 190

gtg ttc cgg gag aag gtc act gag cag cac cgg cag atg ggc aag ggt      741
Val Phe Arg Glu Lys Val Thr Glu Gln His Arg Gln Met Gly Lys Gly
        195                 200                 205

ggc aag cat cac ctt ggc ctg gag gag ccc aag aag ctg cga cca ccc      789
Gly Lys His His Leu Gly Leu Glu Glu Pro Lys Lys Leu Arg Pro Pro
    210                 215                 220

cct gcc agg act ccc tgc caa cag gaa ctg gac cag gtc ctg gag cgg      837
Pro Ala Arg Thr Pro Cys Gln Gln Glu Leu Asp Gln Val Leu Glu Arg
225                 230                 235                 240

atc tcc acc atg cgc ctt ccg gat gag cgg ggc cct ctg gag cac ctc      885
Ile Ser Thr Met Arg Leu Pro Asp Glu Arg Gly Pro Leu Glu His Leu
                245                 250                 255

tac tcc ctg cac atc ccc aac tgt gac aag cat ggc ctg tac aac ctc      933
Tyr Ser Leu His Ile Pro Asn Cys Asp Lys His Gly Leu Tyr Asn Leu
            260                 265                 270

aaa cag tgc aag atg tct ctg aac ggg cag cgt ggg gag tgc tgg tgt      981
Lys Gln Cys Lys Met Ser Leu Asn Gly Gln Arg Gly Glu Cys Trp Cys
        275                 280                 285

gtg aac ccc aac acc ggg aag ctg atc cag gga gcc ccc acc atc cgg     1029
Val Asn Pro Asn Thr Gly Lys Leu Ile Gln Gly Ala Pro Thr Ile Arg
    290                 295                 300

ggg gac ccc gag tgt cat ctc ttc tac aat gag cag cag gag gct tgc     1077
Gly Asp Pro Glu Cys His Leu Phe Tyr Asn Glu Gln Gln Glu Ala Cys
305                 310                 315                 320

ggg gtg cac acc cag cgg atg cag tag accgcagcca gccggtgcct           1124
Gly Val His Thr Gln Arg Met Gln
                325

ggcgcccctg ccccccgccc ctctccaaac accggcagaa aacggagagt gcttgggtgg   1184

tgggtgctgg aggattttcc agttctgaca cacgtattta tatttggaaa gagaccagca   1244

ccgagctcgg cacctccccg gcctctctct tcccagctgc agatgccaca cctgctcctt   1304

cttgctttcc ccgggggagg aaggggttg tggtcgggga gctggggtac aggtttgggg   1364
```

aggggaaga gaaattttta tttttgaacc cctgtgtccc ttttgcataa gattaaagga 1424

aggaaaagt 1433

<210> 4
<211> 328
<212> PRT
<213> Homo sapiens

<400> 4

Met Leu Pro Arg Val Gly Cys Pro Ala Leu Pro Leu Pro Pro Pro Pro
1 5 10 15

Leu Leu Pro Leu Leu Pro Leu Leu Leu Leu Leu Leu Gly Ala Ser Gly
20 25 30

Gly Gly Gly Gly Ala Arg Ala Glu Val Leu Phe Arg Cys Pro Pro Cys
35 40 45

Thr Pro Glu Arg Leu Ala Ala Cys Gly Pro Pro Pro Val Ala Pro Pro
50 55 60

Ala Ala Val Ala Ala Val Ala Gly Gly Ala Arg Met Pro Cys Ala Glu
65 70 75 80

Leu Val Arg Glu Pro Gly Cys Gly Cys Cys Ser Val Cys Ala Arg Leu
85 90 95

Glu Gly Glu Ala Cys Gly Val Tyr Thr Pro Arg Cys Gly Gln Gly Leu
100 105 110

Arg Cys Tyr Pro His Pro Gly Ser Glu Leu Pro Leu Gln Ala Leu Val
115 120 125

Met Gly Glu Gly Thr Cys Glu Lys Arg Arg Asp Ala Glu Tyr Gly Ala
130 135 140

Ser Pro Glu Gln Val Ala Asp Asn Gly Asp Asp His Ser Glu Gly Gly
145 150 155 160

Leu Val Glu Asn His Val Asp Ser Thr Met Asn Met Leu Gly Gly Gly
165 170 175

Gly Ser Ala Gly Arg Lys Pro Leu Lys Ser Gly Met Lys Glu Leu Ala
180 185 190

Val Phe Arg Glu Lys Val Thr Glu Gln His Arg Gln Met Gly Lys Gly

195 200 205

Gly Lys His His Leu Gly Leu Glu Glu Pro Lys Lys Leu Arg Pro Pro
210 215 220

Pro Ala Arg Thr Pro Cys Gln Gln Glu Leu Asp Gln Val Leu Glu Arg
225 230 235 240

Ile Ser Thr Met Arg Leu Pro Asp Glu Arg Gly Pro Leu Glu His Leu
245 250 255

Tyr Ser Leu His Ile Pro Asn Cys Asp Lys His Gly Leu Tyr Asn Leu
260 265 270

Lys Gln Cys Lys Met Ser Leu Asn Gly Gln Arg Gly Glu Cys Trp Cys
275 280 285

Val Asn Pro Asn Thr Gly Lys Leu Ile Gln Gly Ala Pro Thr Ile Arg
290 295 300

Gly Asp Pro Glu Cys His Leu Phe Tyr Asn Glu Gln Gln Glu Ala Cys
305 310 315 320

Gly Val His Thr Gln Arg Met Gln
325

<210> 5
<211> 2497
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (108)..(983)
<223>

<400> 5
ggcacgaggg ccgcttcctg cctggattcc acagcttcgc gccgtgtact gtcgccccat 60

ccctgcgcgc ccagcctgcc aagcagcgtg ccccggttgc aggcgtc atg cag cgg 116
Met Gln Arg
1

gcg cga ccc acg ctc tgg gcc gct gcg ctg act ctg ctg gtg ctg ctc 164
Ala Arg Pro Thr Leu Trp Ala Ala Ala Leu Thr Leu Leu Val Leu Leu
5 10 15

cgc ggg ccg ccg gtg gcg cgg gct ggc gcg agc tcg ggg ggc ttg ggt 212
Arg Gly Pro Pro Val Ala Arg Ala Gly Ala Ser Ser Gly Gly Leu Gly
20 25 30 35

ccc gtg gtg cgc tgc gag ccg tgc gac gcg cgt gca ctg gcc cag tgc 260

```
Pro Val Val Arg Cys Glu Pro Cys Asp Ala Arg Ala Leu Ala Gln Cys
                40                  45                  50

gcg cct ccg ccc gcc gtg tgc gcg gag ctg gtg cgc gag ccg ggc tgc      308
Ala Pro Pro Pro Ala Val Cys Ala Glu Leu Val Arg Glu Pro Gly Cys
            55                  60                  65

ggc tgc tgc ctg acg tgc gca ctg agc gag ggc cag ccg tgc ggc atc      356
Gly Cys Cys Leu Thr Cys Ala Leu Ser Glu Gly Gln Pro Cys Gly Ile
        70                  75                  80

tac acc gag cgc tgt ggc tcc ggc ctt cgc tgc cag ccg tcg ccc gac      404
Tyr Thr Glu Arg Cys Gly Ser Gly Leu Arg Cys Gln Pro Ser Pro Asp
    85                  90                  95

gag gcg cga ccg ctg cag gcg ctg ctg gac ggc cgc ggg ctc tgc gtc      452
Glu Ala Arg Pro Leu Gln Ala Leu Leu Asp Gly Arg Gly Leu Cys Val
100                 105                 110                 115

aac gct agt gcc gtc agc cgc ctg cgc gcc tac ctg ctg cca gcg ccg      500
Asn Ala Ser Ala Val Ser Arg Leu Arg Ala Tyr Leu Leu Pro Ala Pro
                120                 125                 130

cca gct cca gga aat gct agt gag tcg gag gaa gac cgc agc gcc ggc      548
Pro Ala Pro Gly Asn Ala Ser Glu Ser Glu Glu Asp Arg Ser Ala Gly
            135                 140                 145

agt gtg gag agc ccg tcc gtc tcc agc acg cac cgg gtg tct gat ccc      596
Ser Val Glu Ser Pro Ser Val Ser Ser Thr His Arg Val Ser Asp Pro
        150                 155                 160

aag ttc cac ccc ctc cat tca aag ata atc atc atc aag aaa ggg cat      644
Lys Phe His Pro Leu His Ser Lys Ile Ile Ile Ile Lys Lys Gly His
    165                 170                 175

gct aaa gac agc cag cgc tac aaa gtt gac tac gag tct cag agc aca      692
Ala Lys Asp Ser Gln Arg Tyr Lys Val Asp Tyr Glu Ser Gln Ser Thr
180                 185                 190                 195

gat acc cag aac ttc tcc tcc gag tcc aag cgg gag aca gaa tat ggt      740
Asp Thr Gln Asn Phe Ser Ser Glu Ser Lys Arg Glu Thr Glu Tyr Gly
                200                 205                 210

ccc tgc cgt aga gaa atg gaa gac aca ctg aat cac ctg aag ttc ctc      788
Pro Cys Arg Arg Glu Met Glu Asp Thr Leu Asn His Leu Lys Phe Leu
            215                 220                 225

aat gtg ctg agt ccc agg ggt gta cac att ccc aac tgt gac aag aag      836
Asn Val Leu Ser Pro Arg Gly Val His Ile Pro Asn Cys Asp Lys Lys
        230                 235                 240

gga ttt tat aag aaa aag cag tgt cgc cct tcc aaa ggc agg aag cgg      884
Gly Phe Tyr Lys Lys Lys Gln Cys Arg Pro Ser Lys Gly Arg Lys Arg
    245                 250                 255

ggc ttc tgc tgg tgt gtg gat aag tat ggg cag cct ctc cca ggc tac      932
Gly Phe Cys Trp Cys Val Asp Lys Tyr Gly Gln Pro Leu Pro Gly Tyr
260                 265                 270                 275

acc acc aag ggg aag gag gac gtg cac tgc tac agc atg cag agc aag      980
Thr Thr Lys Gly Lys Glu Asp Val His Cys Tyr Ser Met Gln Ser Lys
                280                 285                 290
```

tag acgcctgccg caaggttaat gtggagctca aatatgcctt attttgcaca 1033

aaagactgcc aaggacatga ccagcagctg gctacagcct cgatttatat ttctgtttgt 1093

ggtgaactga ttttttttta aaccaaagtt tagaaagagg tttttgaaat gcctatggtt 1153

tctttgaatg gtaaacttga gcatcttttc actttccagt agtcagcaaa gagcagtttg 1213

aattttcttg tcgcttccta tcaaaatatt cagagactcg agcacagcac ccatacttca 1273

tgcgcccgtg gaatgctcac cacatgttgg tcgaagcggc cgaccactga ctttgtgact 1333

taggcggctg tgttgcctat gtagagaaca cgcttcaccc ccactccccg tacagtgcgc 1393

acaggcttta tcgagaatag gaaaaccttt aaaccccggt catccggaca tcccaacgca 1453

tgctcctgga gctcacagcc ttctgtggtg tcatttctga aacaagggcg tggatccctc 1513

aaccaagaag aatgtttatg tcttcaagtg acctgtactg cttggggact attggagaaa 1573

ataaggtgga gtcctacttg tttaaaaaat atgtatctaa gaatgttcta gggcactctg 1633

ggaacctata aaggcaggta tttcgggccc tcctcttcag gaatcttcct gaagacatgg 1693

cccagtcgaa ggcccaggat ggcttttgct gcggccccgt ggggtaggag ggacagagag 1753

acagggagag tcagcctcca cattcagagg catcacaagt aatggcacaa ttcttcggat 1813

gactgcagaa aatagtgttt tgtagttcaa caactcaaga cgaagcttat ttctgaggat 1873

aagctcttta aaggcaaagc tttattttca tctctcatct tttgtcctcc ttagcacaat 1933

gtaaaaaaga atagtaatat cagaacagga aggaggaatg gcttgctggg gagcccatcc 1993

aggacactgg gagcacatag agattcaccc atgtttgttg aacttagagt cattctcatg 2053

cttttcttta taattcacac atatatgcag agaagatatg ttcttgttaa cattgtatac 2113

aacatagccc caaatatagt aagatctata ctagataatc ctagatgaaa tgttagagat 2173

gctatatgat acaactgtgg ccatgactga ggaaaggagc tcacgcccag agactgggct 2233

gctctcccgg aggccaaacc caagaaggtc tggcaaagtc aggctcaggg agactctgcc 2293

ctgctgcaga cctcggtgtg gacacacgct gcatagagct ctccttgaaa acagaggggt 2353

ctcaagacat tctgcctacc tattagcttt tctttatttt tttaactttt tgggggggaaa 2413

agtatttttg agaagtttgt cttgcaatgt atttataaat agtaaataaa gtttttacca 2473

ttaaaaaaaa aaaaaaaaaa aaaa 2497

<210> 6
<211> 291
<212> PRT
<213> Homo sapiens

<400> 6

Met Gln Arg Ala Arg Pro Thr Leu Trp Ala Ala Ala Leu Thr Leu Leu

```
1               5                  10                  15

Val Leu Leu Arg Gly Pro Pro Val Ala Arg Ala Gly Ala Ser Ser Gly
            20                  25                  30

Gly Leu Gly Pro Val Val Arg Cys Glu Pro Cys Asp Ala Arg Ala Leu
        35                  40                  45

Ala Gln Cys Ala Pro Pro Pro Ala Val Cys Ala Glu Leu Val Arg Glu
    50                  55                  60

Pro Gly Cys Gly Cys Cys Leu Thr Cys Ala Leu Ser Glu Gly Gln Pro
65                  70                  75                  80

Cys Gly Ile Tyr Thr Glu Arg Cys Gly Ser Gly Leu Arg Cys Gln Pro
                85                  90                  95

Ser Pro Asp Glu Ala Arg Pro Leu Gln Ala Leu Leu Asp Gly Arg Gly
            100                 105                 110

Leu Cys Val Asn Ala Ser Ala Val Ser Arg Leu Arg Ala Tyr Leu Leu
        115                 120                 125

Pro Ala Pro Pro Ala Pro Gly Asn Ala Ser Glu Ser Glu Glu Asp Arg
    130                 135                 140

Ser Ala Gly Ser Val Glu Ser Pro Ser Val Ser Ser Thr His Arg Val
145                 150                 155                 160

Ser Asp Pro Lys Phe His Pro Leu His Ser Lys Ile Ile Ile Ile Lys
                165                 170                 175

Lys Gly His Ala Lys Asp Ser Gln Arg Tyr Lys Val Asp Tyr Glu Ser
            180                 185                 190

Gln Ser Thr Asp Thr Gln Asn Phe Ser Ser Glu Ser Lys Arg Glu Thr
        195                 200                 205

Glu Tyr Gly Pro Cys Arg Arg Glu Met Glu Asp Thr Leu Asn His Leu
    210                 215                 220

Lys Phe Leu Asn Val Leu Ser Pro Arg Gly Val His Ile Pro Asn Cys
225                 230                 235                 240

Asp Lys Lys Gly Phe Tyr Lys Lys Lys Gln Cys Arg Pro Ser Lys Gly
                245                 250                 255
```

Arg Lys Arg Gly Phe Cys Trp Cys Val Asp Lys Tyr Gly Gln Pro Leu
260 265 270

Pro Gly Tyr Thr Thr Lys Gly Lys Glu Asp Val His Cys Tyr Ser Met
275 280 285

Gln Ser Lys
290

<210> 7
<211> 6128
<212> DNA
<213> Homo sapiens

<220>
<221> mRNA
<222> (597)..(1110)
<223> Exon 1

<220>
<221> mRNA
<222> (2657)..(2826)
<223> Exon 2

<220>
<221> mRNA
<222> (4041)..(4169)
<223> Exon 3

<220>
<221> mRNA
<222> (5069)..(5769)
<223> Exon 4

<220>
<221> CDS
<222> (762)..(1110)
<223>

<220>
<221> CDS
<222> (2657)..(2826)
<223>

<220>
<221> CDS
<222> (4041)..(4169)
<223>

<220>
<221> CDS

<222> (5069)..(5197)
<223>

<400> 7
ggggagcttg gcagagccct aggatgaact gtgttttgtt ttgttttgtt ttgttgcagg 60

ttggagcctt tttcctgcag aaagagaagc aattccgggc ttttccacag ataacctgct 120

agggctcgca gagacctgga ttctaacccc atctcgcctt tcctcacctg gggcattgtt 180

ttctgcattt gagaactgct gggtgggctg gtgttaccca ggatttttaa atttttgcag 240

gagacgcttt gcaggagatc tctagggctt ttttctgtgt caattaaaga gcctggccag 300

ctggacctgg gctgtctttt tgacaaaaac aaacgtcatc cccctcccag ctgagcactt 360

gttagaactg gactttaact gagggcctga acccccctaac aacgggacaa acagtatgaa 420

atcaaaaccg tttaccctcc tcccaccagc ggtttgcgta gggccttggg tgcactagca 480

aaacaaactt attttgaaca ctcagctcct agcgtgcggc gctgccaatc attaacctcc 540

tggtgcaagt ggcgcggcct gtgcccttta taaggtgcgc gctgtgtcca gcgagcatcg 600

gccaccgcca tcccatccag cgagcatctg ccgccgcgcc gccgccaccc tcccagagag 660

cactggccac cgctccacca tcacttgccc agagtttggg ccaccgcccg ccgccaccag 720

cccagagagc atcggcccct gtctgctgct cgcgcctgga g atg tca gag gtc ccc 776
Met Ser Glu Val Pro
1 5

gtt gct cgc gtc tgg ctg gta ctg ctc ctg ctg act gtc cag gtc ggc 824
Val Ala Arg Val Trp Leu Val Leu Leu Leu Leu Thr Val Gln Val Gly
10 15 20

gtg aca gcc ggc gct ccg tgg cag tgc gcg ccc tgc tcc gcc gag aag 872
Val Thr Ala Gly Ala Pro Trp Gln Cys Ala Pro Cys Ser Ala Glu Lys
25 30 35

ctc gcg ctc tgc ccg ccg gtg tcc gcc tcg tgc tcg gag gtc acc cgg 920
Leu Ala Leu Cys Pro Pro Val Ser Ala Ser Cys Ser Glu Val Thr Arg
40 45 50

tcc gcc ggc tgc ggc tgt tgc ccg atg tgc gcc ctg cct ctg ggc gcc 968
Ser Ala Gly Cys Gly Cys Cys Pro Met Cys Ala Leu Pro Leu Gly Ala
55 60 65

gcg tgc ggc gtg gcg act gca cgc tgc gcc cgg gga ctc agt tgc cgc 1016
Ala Cys Gly Val Ala Thr Ala Arg Cys Ala Arg Gly Leu Ser Cys Arg
70 75 80 85

gcg ctg ccg ggg gag cag caa cct ctg cac gcc ctc acc cgc ggc caa 1064
Ala Leu Pro Gly Glu Gln Gln Pro Leu His Ala Leu Thr Arg Gly Gln
90 95 100

ggc gcc tgc gtg cag gag tct gac gcc tcc gct ccc cat gct gca g 1110
Gly Ala Cys Val Gln Glu Ser Asp Ala Ser Ala Pro His Ala Ala
105 110 115

gtaccacagt cccgcccctg ctccagcacc tcctgccgcc cgcccccgcc cggcacctcc 1170

```
cgcccccact cccctaacta ctgggtgggg ctgcagccgg gcaggggctt gtccacaact   1230

agagcttgaa accagagcac gtagttgggg aaggagcttg ggtcacccag tggagcccgc   1290

tcattgcacg gtcttggcag gacgtgccct gggagaagaa ggaagatgtt ccagggcaca   1350

catagcttag tggagactca aggaggaagc tcccacttgg ccctcgtagc ccagggatct   1410

ttagagaccc agacagcaag agtagaggga ggcggtgggg ctgcctccgc agttgggatg   1470

ccaccctggg tcccccacca gtctccttgc cagcacgtgg gcagccccag tggctggagc   1530

atccgggaag gcctaagaag ggaatccctc ctgctggcct ccccaggagg tgtttggaat   1590

gtcctttgta atatacgtcc tggatacagt atgtgcttcc cagatgttta cagaacataa   1650

tgtgagagtt gaggctaaaa gcttcacttc cattcattgc catagaaaaa caatgtttga   1710

tgtgtatgtt gccccctaag aacatgacaa tcaggttctt tcctggattt aagcaggcat   1770

ttgtttgtag cgggaagtgg aggaaatgcc tctttctcta ctccagcccc tatctcctca   1830

tctggactgc ttgcacgtcc tgtcctgtgc cattcttgga ggaattagga cttccatacc   1890

ctgatagggt gttagagtga gtgtgtgaga gtcagtgacc cttccaagga gggaacaggg   1950

atttcctgcg gattctctgc ccttcagagg aaatgaaagg gagtcgtctc actgatcctc   2010

aaggagatgg atttccacag ggggtcccct ataggccagt gtgctgattc tgttcatgta   2070

cctttggact ctatctgcaa gatcctctcg ccatgggggt ggagggggtg ttgcagagaa   2130

gtccattagg acgtgggaag gcaggggccc caaaatagaa gagggtagta gtttgggtca   2190

ggagagctgc attctagact ccacttctag atgaactgta ccatccattc agggatatta   2250

gtaaaatccc cattttaagg atggagaagc tgaggtttcg aggctcccta ctaatgcggt   2310

taggaactga actctaatcc tggcttcttt agggttgttc tcaacctcag caaatgcatt   2370

tgttgctctg ctccctgct gctttaaaag cagaaacgtg gggcaccctt gccaggccct   2430

gaaacttact attgtcagga acagagctgg catgctgttg ggacataact cttcctggaa   2490

accattacaa ggtcctttac agtgagagtc agtgaaaatc aggttaggga atgtggccat   2550

cctcatggcc cagcctggtt ttgagggaag gaagtgatct cctgggggcc cgaaaggggc   2610

agatgctttg ggagggctca tggggtctgc aatgtttcct ttccag ag  gca ggg      2664
                                                    Glu Ala Gly

agc cct gaa agc cca gag agc acg gag ata act gag gag gag ctc ctg      2712
Ser Pro Glu Ser Pro Glu Ser Thr Glu Ile Thr Glu Glu Glu Leu Leu
120                 125                 130                 135

gat aat ttc cat ctg atg gcc cct tct gaa gag gat cat tcc atc ctt      2760
Asp Asn Phe His Leu Met Ala Pro Ser Glu Glu Asp His Ser Ile Leu
                140                 145                 150

tgg gac gcc atc agt acc tat gat ggc tcg aag gct ctc cat gtc acc      2808
```

```
Trp Asp Ala Ile Ser Thr Tyr Asp Gly Ser Lys Ala Leu His Val Thr
            155                 160                 165

aac atc aaa aaa tgg aag gtgaggccca gcaggtgggt ggtgggaact          2856
Asn Ile Lys Lys Trp Lys
        170

ctcctgtcag aggatgtagc ttgagtcggc tgtgacaagc agacctggcc cactccttcc  2916

tggtcctgat gccctggaaa gaaatgcttt tcccccaaaa cctagtggcc atttctcagt  2976

taaactttgt caaactgtga acacatggac tttcaggaga ggagcccact ggatgcaaag  3036

aaaatgtagt tccggacgat aggttacttc atgtctctga gctcctctct aaagggtgtg  3096

agaacattga gttacattgt tccaaaactt agtacaaggg cctaactgct ttcaccccca  3156

ttgtatgatt tgcctccctt taaaaaagtt gtcataaaat agacataaaa tgtactgttt  3216

taactatttt gaagcatagg tttctgtggc attaagtaac aataagtaat gcatgttatt  3276

ctgcaaccac taccaatatc taccttcaga attttcatct tcccaaactg aagctcggaa  3336

ctcattaaac atcaactctc cattatctta attatgtctg ttttgaattc taccacttga  3396

gggcgacaga gaatcaaagt cggagtgatt tgatttgaga aaataaatta cttccacaaa  3456

catgcctacc cagaaaaagg gaggattaaa ccattttctc aacattttgg aatttcatct  3516

gagaataacg ataaatggag gaaagttaat tcttaaatta tccaaatgaa tgccttaatc  3576

ctcacttgat ttagtttatg taaaagaacc ttgtagattg cagaggaaag ttattcaaaa  3636

tctagaaatg tcttctgcta aattcttaat taagagttat gcccactctt ctaaccagat  3696

ggcgggtgcc aaaagtattc agtaggtggc aggaagagtt gggaagagtt ttgtggtgtt  3756

ggtgaggctc ttgccagggg atgtctctgt gcaggtgttg agagagtgat tggtgaggga  3816

gagaatggtg ttcttgttct ttatctgaac aacttaagac agactaactg caggtgggaa  3876

aaagaggccc agagggagca catcacatga cacctagtcc caggtccccg tggccagatc  3936

gtagggccac tcctgcttct acaaaacctc ttccatcagt cattgtctag gctgataatt  3996

cactgtgtct tgttagatat gctaatgtca agttattctc ttag gag ccc tgc cga    4052
                                                Glu Pro Cys Arg
                                                    175

ata gaa ctc tac aga gtc gta gag agt tta gcc aag gca cag gag aca     4100
Ile Glu Leu Tyr Arg Val Val Glu Ser Leu Ala Lys Ala Gln Glu Thr
        180                 185                 190

tca gga gaa gaa att tcc aaa ttt tac ctg cca aac tgc aac aag aat     4148
Ser Gly Glu Glu Ile Ser Lys Phe Tyr Leu Pro Asn Cys Asn Lys Asn
    195                 200                 205

gga ttt tat cac agc aga cag gtaggtggcc ttgccagtgt gcgtcgtcag       4199
Gly Phe Tyr His Ser Arg Gln
210                 215

ggtgaaaggg actactgccc tacattcctg ccaagccacg gtcattcatg tcaaagaagg  4259
```

```
tccactccaa agtagacacc agaagtggtt gtattgagcc agatccaccc ctctgggaac   4319

ctggaacagc taggtgaaga aagcctattg aacaggtcaa tatgtcacca ccagaggtgg   4379

aaaacctgcg ctggaccagg ggccctgggg ctaagcctga gctccatcac tcagcaagcc   4439

aatagctatg gaagactttc ctaggccccc taacgcttgg atcttgggtt cctcacctat   4499

acaagagagg aagggaacaa acttaactta gaaaagcaag tatctatttg gacagctctt   4559

aacacttttt tctccttgaa ctgtctaata aaacacaaca acatttttgt ggcttgccag   4619

ataataaatg agatcaaaca gaccaagatt tcatattgag gagtgcttta ggtctcagtg   4679

aagtacaggt tctgtagatt ttattgggag aaactgagga ctaggccctg cttcacgggc   4739

aatgaacagt ggggcacaca cgagacatgt tccctctggg atgggctccc ctgacatcag   4799

gctatgaagc agacagctgt acacacactg tactgtttaa cacacatggg aagtcattat   4859

tgcacatgcc actcggtcac tgtctctatt ttatgatgag aaccaggagg gtgtgagatt   4919

tgcctgcatc atgggcagct ggtttcacag ccgggactct tggcttggct ctgcagtgct   4979

cggctgcact gaaaaaaggg ccagctatgg ctctactttc cctgtcagct tgtcatctgc   5039

tgctcttgac cttggtcttg tctttgcag tgt gag aca tcc atg gat gga gag      5092
                                Cys Glu Thr Ser Met Asp Gly Glu
                                220

gcg gga ctc tgc tgg tgc gtc tac cct tgg aat ggg aag agg atc cct      5140
Ala Gly Leu Cys Trp Cys Val Tyr Pro Trp Asn Gly Lys Arg Ile Pro
225                 230                 235                 240

ggg tct cca gag atc agg gga gac ccc aac tgc cag atg tat ttt aat      5188
Gly Ser Pro Glu Ile Arg Gly Asp Pro Asn Cys Gln Met Tyr Phe Asn
                245                 250                 255

gta caa aac tgaaaccaga tgaaataatg ttctgtcacg tgaaatattt              5237
Val Gln Asn

aagtatatag tatatttata ctctagaaca tgcacattta tatatatatg tatatgtata   5297

tatatatagt aactactttt tatactccat acataacttg atatagaaag ctgtttattt   5357

attcactgta agtttatttt ttctacacag taaaaacttg tactatgtta ataacttgtc   5417

ctatgtcaat ttgtatatca tgaaacactt ctcatcatat tgtatgtaag taattgcatt   5477

tctgctcttc caaagctcct gcgtctgttt ttaaagagca tggaaaaata ctgcctagaa   5537

aatgcaaaat gaaataagag agagtagttt ttcagctagt ttgaaggagg acggttaact   5597

tgtatattcc accattcaca tttgatgtac atgtgtaggg aaagttaaaa gtgttgatta   5657

cataatcaaa gctacctgtg gtgatgttgc cacctgttaa aatgtacact ggatatgttg   5717

ttaaacacgt gtctataatg gaaacattta caataaatat tctgcatgga aatactgtta   5777

tctctgtgga tgtttattta gtgatattat ctttttgaaa tcccttttgtc cttcaccctc   5837
```

accttcctta tgcattttgt gcttttgtga taccaagatg gtgttttggg tctgcagagc 5897

cccttgtcct aactacccta gatcttgccc cggctgcccc agggtttcca cttcccagtg 5957

gcattgggac cttcctcctc tgcctggcca ggctgggcat tcctgtgtcc cacctctcca 6017

ttagtcccca tcctccccag ggtcctcagc ctcttcctct tggtttggtt tgtcccttcc 6077

gttcaactta taaataggaa cagggtttct ttcctcttaa acacaagaca g 6128

<210> 8
<211> 259
<212> PRT
<213> Homo sapiens

<400> 8

Met Ser Glu Val Pro Val Ala Arg Val Trp Leu Val Leu Leu Leu Leu
1 5 10 15

Thr Val Gln Val Gly Val Thr Ala Gly Ala Pro Trp Gln Cys Ala Pro
20 25 30

Cys Ser Ala Glu Lys Leu Ala Leu Cys Pro Pro Val Ser Ala Ser Cys
35 40 45

Ser Glu Val Thr Arg Ser Ala Gly Cys Gly Cys Cys Pro Met Cys Ala
50 55 60

Leu Pro Leu Gly Ala Ala Cys Gly Val Ala Thr Ala Arg Cys Ala Arg
65 70 75 80

Gly Leu Ser Cys Arg Ala Leu Pro Gly Glu Gln Gln Pro Leu His Ala
85 90 95

Leu Thr Arg Gly Gln Gly Ala Cys Val Gln Glu Ser Asp Ala Ser Ala
100 105 110

Pro His Ala Ala Glu Ala Gly Ser Pro Glu Ser Pro Glu Ser Thr Glu
115 120 125

Ile Thr Glu Glu Glu Leu Leu Asp Asn Phe His Leu Met Ala Pro Ser
130 135 140

Glu Glu Asp His Ser Ile Leu Trp Asp Ala Ile Ser Thr Tyr Asp Gly
145 150 155 160

Ser Lys Ala Leu His Val Thr Asn Ile Lys Lys Trp Lys Glu Pro Cys
165 170 175

Arg Ile Glu Leu Tyr Arg Val Val Glu Ser Leu Ala Lys Ala Gln Glu
180 185 190

Thr Ser Gly Glu Glu Ile Ser Lys Phe Tyr Leu Pro Asn Cys Asn Lys
195 200 205

Asn Gly Phe Tyr His Ser Arg Gln Cys Glu Thr Ser Met Asp Gly Glu
210 215 220

Ala Gly Leu Cys Trp Cys Val Tyr Pro Trp Asn Gly Lys Arg Ile Pro
225 230 235 240

Gly Ser Pro Glu Ile Arg Gly Asp Pro Asn Cys Gln Met Tyr Phe Asn
245 250 255

Val Gln Asn

<210> 9
<211> 31024
<212> DNA
<213> Homo sapiens

<220>
<221> gene
<222> (1)..(31024)
<223>

<220>
<221> mRNA
<222> (1)..(556)
<223> Exon 1

<220>
<221> mRNA
<222> (27148)..(27377)
<223> Exon 2

<220>
<221> mRNA
<222> (28449)..(28589)
<223> Exon 3

<220>
<221> mRNA
<222> (30531)..(31024)
<223> Exon 4

<220>
<221> CDS

<222> (115)..(556)
<223>

<220>
<221> CDS
<222> (27148)..(27377)
<223>

<220>
<221> CDS
<222> (28449)..(28589)
<223>

<220>
<221> CDS
<222> (30531)..(30692)
<223>

<400> 9
ggcgagggag gaggaagaag cggaggaggc ggctcccgcg ctcgcagggc cgtgccacct 60

gcccgcccgc ccgctcgctc gctcgcccgc cgcgccgcgc tgccgaccgc cagc atg 117
Met
1

ctg ccg aga gtg ggc tgc ccc gcg ctg ccg ctg ccg ccg ccg ccg ctg 165
Leu Pro Arg Val Gly Cys Pro Ala Leu Pro Leu Pro Pro Pro Pro Leu
5 10 15

ctg ccg ctg ctg ctg ctg cta ctg ggc gcg agt ggc ggc ggc ggc ggg 213
Leu Pro Leu Leu Leu Leu Leu Leu Gly Ala Ser Gly Gly Gly Gly Gly
20 25 30

gcg cgc gcg gag gtg ctg ttc cgc tgc ccg ccc tgc aca ccc gag cgc 261
Ala Arg Ala Glu Val Leu Phe Arg Cys Pro Pro Cys Thr Pro Glu Arg
35 40 45

ctg gcc gcc tgc ggg ccc ccg ccg gtt gcg ccg ccc gcc gcg gtg gcc 309
Leu Ala Ala Cys Gly Pro Pro Pro Val Ala Pro Pro Ala Ala Val Ala
50 55 60 65

gca gtg gcc gga ggc gcc cgc atg cca tgc gcg gag ctc gtc cgg gag 357
Ala Val Ala Gly Gly Ala Arg Met Pro Cys Ala Glu Leu Val Arg Glu
70 75 80

ccg ggc tgc ggc tgc tgc tcg gtg tgc gcc cgg ctg gag ggc gag gcg 405
Pro Gly Cys Gly Cys Cys Ser Val Cys Ala Arg Leu Glu Gly Glu Ala
85 90 95

tgc ggc gtc tac acc ccg cgc tgc ggc cag ggg ctg cgc tgc tat ccc 453
Cys Gly Val Tyr Thr Pro Arg Cys Gly Gln Gly Leu Arg Cys Tyr Pro
100 105 110

cac ccg ggc tcc gag ctg ccc ctg cag gcg ctg gtc atg ggc gag ggc 501
His Pro Gly Ser Glu Leu Pro Leu Gln Ala Leu Val Met Gly Glu Gly
115 120 125

act tgt gag aag cgc cgg gac gcc gag tat ggc gcc agc ccg gag cag 549

Thr Cys Glu Lys Arg Arg Asp Ala Glu Tyr Gly Ala Ser Pro Glu Gln
130 135 140 145

gtt gca g gtaacgcggt ctggaacaag tagttgggag aaacttggag ggcagcggag 606
Val Ala

aagcccgacg ggcggctgga ccttacggac ggttttaggg gcggcagagc cgagaccttg 666

gaccaaatca aggggactg ttgctagcgg gacgcggaag tcaggcccgg ggaggggaac 726

tggagttaga gcaagtggaa gagccctggc gactcatttg gtggggatga atggggagag 786

gaaaggcatc cctgctgcca tagttttacc cgcgtggaac ggacccagac ttctattctg 846

gaggagggat ttgctgagaa gggggccggt ggtcagatcc atgggagtag gaactcgaat 906

cttatttggg gcactgcttg aggacccaag ggtcctggtg ggatcgctgc tgggaggcgg 966

aggctgcggg gtgccagtga caccaggaac cttggaagag gagagggagc taggaggagg 1026

caccggcgtc ccaaggggcc cttccttcta ctggggttgt gagttggaaa gaaaggactt 1086

ggtactttgt aggatgagtc gcttaggaga aaaatgaaaa ccaccactac tatttctcca 1146

tagggaacga gcagggctct tgctaccacg ttgaatggaa aaataaaagg gaggcatttg 1206

gtaggcgagg ggtctcctag cagctgaatc aatgtaactt aagcccgatc gccggctgca 1266

ttccctggac gggggcagcc aggctgtgcg gtttcgcatc tgggccccag atcgaaaggc 1326

aggagagatg cgttgccacc ccccaccccc gcccctatct taaattgagg gacccttcc 1386

tgtaggatag aatgaggacc cgatattctt tgttcagggg ccctcatccc tcccaaaaag 1446

ttatttgagg gacacaacca ggctctaagg aggttacttt ttttttttt taattacgaa 1506

agcctcctgc cccagtggtc gggaggaagc tgtaattatt ttctttcctc ttggagacag 1566

tggtgaaatg tgtccactgc ctgctgctgt tgttgaaaaa gcaagtgagt cttaagtttc 1626

tcaaacttaa aaagaaaaag tgaaaagaag aatgcttgct ttttagcttc tgaggtcttc 1686

gtaagccccc caccccatgc accaacccgc ctcccttctc ctggtcggct attgtttttt 1746

tccctgttct ccttgagaat ggaattaggg gtcagagttg aaatctttct tctgtttctc 1806

tagcacattt tgctctggag tacagagcaa gttagggctt cctctgagtg gagcctctca 1866

gtccctttct tctccctcgg ctcgcttcct tccctagcct gcctggaagt taacccttcc 1926

ctcctgcttc tcttagaaag taacaataaa aggaaaaagc ccccacgaaa tcctttcact 1986

gcggggtgcc tgtcccccag tcatgcgttc ctcctcatca cccctcgtgg tccaccgggg 2046

aggtgctgtg cagtttctga gctgaactct gacaaactaa acatcttcct cagcagggtg 2106

aacttttga aaggcctcta agttcatagg atggtttggc tgctgactgt tttaataaag 2166

tgtgtgatgc ctttccttct ttaacgtttt actttttttt ttttcctact cttcattcgt 2226

cttattctaa gcctgaaaac caattctact tcccaaatcc ctgaagccgg ttagggtctt 2286

tagccaggaa cagatgcggc agattttctc ccattcactt ttccctggag cctgaagggg 2346

gtgggggta gttgtcaggg gaggccaggg acaatggggt gagttcatgg gaagaatgtc 2406

actctggatt ttctgcctgg ttatgggact cccttccgcc agcttgcccc agctgcgttc 2466

tcctccccag tgcccagacg ccttccgcat ttttgtgcac cggcctggga agaggggatg 2526

ttaagcaggg gtggaaggtg gggggtgggg ctggaagctg agtgggtgca ctcaaaagaa 2586

gtgaccaatg tcaggggtgc ctgagaaaga gtcgggctgg gggaggggag cactttggga 2646

aggagcctga gcttgcaagg gtgcaggaga aaaaggggtc aaggttggag gcttttgtgt 2706

gggagagggt cttaggctgg ggcagagact ggagtcactg gatggagagt gctgggtccg 2766

agttgggtgg ggcctggccc agacggagat cttagcactg aatcctctgc ctcccactcc 2826

atctcccgtc ccgtcttctc tcctctccga agaatgagca ccccccctcc agcctgcatt 2886

cctctttgag caggttgctt ttgagaattt taaacatacc ctcttatttc gaggcagcga 2946

acctcagcag ttccggcttt gctgttagct cttttgtctt tgttcccggg actcaggggc 3006

tgtgttgctg ctggcttcct ggggatggca aggatgcagg ggtgggggtg ggcacatcag 3066

tgggtggaag gagtgggtgg tggtgagcct gtatcccacg gggacgatgg gctggggctt 3126

gttggttcaa agaaactgag aagtcagtct tgggggaagg ggtgggtagg cctagaagaa 3186

aaaccgaaag gagaccctga gatctgcttg aagttgggac tctaaaagcg tgtgtatgtg 3246

tttggtggtg gagtggatag aagtgcctag aaatgtcagg gtcgaggagg gaacagggtt 3306

gggccagtcc aggccctttc agaattaagg ctgggctaga ccggttcctt cagtatctcc 3366

ccagcctcct tccctaccgt gtacaatgca ggggattatg ctgaggggcg gggccgcccg 3426

agtggccgtt tggcccgctg cacgtataaa ggtgctacag aaatgtgtag tcgttcaaag 3486

tcccagggcg ggcagacagg caggcaggca ggcgggcggg cgggcgggcg ggcgggcctg 3546

gggaggctct gcatctgcag ttgccgggga agaataaagg ggggaagagg caagcactga 3606

tttacagtgg ttcagtgcag cttcaaaaaa aaaggggaa aaagcaatgg gacttctggc 3666

tcagagtact gctttgcatc tcgtcttgca taaatttgca tactgagttc aaagtcgtaa 3726

aatagactag ggaaatgagt gggatctact gtcactgatt gttaacacca ctaccttcca 3786

cctcccctcc ccccatctgg acagaactcc cggcactagt tctgctaagc agcgctgcct 3846

gagtcaccag tcaccactgt tcatgaaaga aggagatggg gaagcgcctg tctaggagac 3906

caagcagggt gtggcagtgt gcaccctggc ctcggaggca gggcctctct gaagggtgtg 3966

gcccgacaga ttctggacca gtgaggcaag tggtgttctt tctgacgcct ggggctgccg 4026

ttgcccatgg tgggctgcag tttgaggaag ctgcagaatc cgggtcttct ttgtttgaaa 4086

tttccctact cagcttccac atatgtccaa aacatttact gtgataagtc cctttgtgtg 4146

gtggtaatga cccacttctt tagacaagat atttgaggaa ggttcgagaa cagaaaaatt 4206

ctacttaaca atgttaagga atctgagaaa agtggggcac aggagtctct aggttagaat 4266

aggtggtaga gcctggaagg agagaagaac ccatctgttg tattcccgag ggtctgccca 4326

cacaatggct cttggccaag tggactgttg agcttcctgt caggcagagg gaagagggag 4386

gcttagatag ttctcactgt tggcctactt cagcactgag cttcaagact ttcaggcaaa 4446

tcattccttc aggctcagag gctggaaatt agttctataa aaatgccatg gggccggtcg 4506

cattggctca tgcctgtaat cccagcactt tgggaggccg aggcaggtgg atcacaaggt 4566

caggagtttg gagaccagcc tggccagtat agtgaaacct cttctctgct acaaatacaa 4626

aaattagctg ggcatggtgg agcatgcctg tagtcccagc tgcttgggag gctgaggcag 4686

gagaattgct tgaacccggg aggtggagct tgcagtgagc tgggattgca tcactgctct 4746

ccagcccggg caacagagca aaaaaactcc gtctcaaaaa ataaataaaa taaaaatgcc 4806

acgggttggc caggcgcggt ggctcacgcc tgtaattcca gcacttaggg aggctgaggt 4866

gggcagatca cctaaggcca ggagttcaag accagcctgg ctaacatagt gaaaccccat 4926

ctctactaaa aatgcaaaaa ttagttgggc atggtggcac ctgcctgtag tcccagctgc 4986

tcgggactct gaggtaggag aatcgcttga acctgggaac tggaggttgc agtgagccga 5046

gatggcgcca ctccactcca gcctgggcaa caagagtgaa actccgtctc aaaaataaaa 5106

taaaataaaa ataaaaataa aaagccatgg gttacctgaa agtaggagca tatccattca 5166

ctcattatct gacctttgag ggctgacttg gagagaactg ttctctggcc cagtcatcag 5226

gtggcttcca gttggttctc agctgggctg ttcacaccca tagaaaatgg gaggtacgac 5286

tgttcaattt ctttttttt ttgagacgga gactcgctct cttgcccagg ctggagtgca 5346

gtggtgcgat ctcggctcac tgcaagctcc gtctcctggg ttcacgccat tctcctgcct 5406

cagcctcccg agtagttggg actacaggca cctgccacca cgcccggcta acttttgta 5466

ttttttagta gagacggggt ttcacagtgt tagccaggat ggtctcgatc tcctgacctc 5526

atgatccgcc cacctcggcc tcccaaagtg ctggggatta caggcatgag ccacggtgcc 5586

cgaccaggct aatttttttt ttttttttt ttttgagac ggagtcttgc tccgtcgccc 5646

aggctggagt gcagtcgccc agcctggagt gcagtggcgt gatcttggct cactgcaagc 5706

tctgcttcct gggttcacgc cattgtcctg cctcagcctc ccgagtagtt gggaattcac 5766

gcatgcacca ccatacacgg ttaatatttt tgtattttta gtagaggcag ggttttacca 5826

tgttggccag gttggtctca aactcgtgaa cctcaagtga tctgcccgcc ttggcctccc 5886

aaagtgctgg gattacaggc atgagctacc gcacctggcc acaactgttc gatattgata 5946

tgaggaagtt ctgaaaattc tgcaccaaag ggaatggaga acaggtaaat acctatgatg 6006

cctggcagga taggacaggg ggctgctgaa aggaatgaag gaggagcaga acaaaaaggg 6066

gctgtgaggc cacatccaca tctgcttttg ttgttggtcc ttcccgcgta ctagaagact 6126

gtacccaaaa gaaaagaaaa aagcttgttt ttttttgtt tgttttttg tttttacaaa 6186

gagtctgtgt atctgggaat tgttattttt atttgtatta ttatttttat tcatttattt 6246

atttcaagac cgagttttgc tctgtcaccc aggctggagt gcagtggcgc aatctcggtt 6306

cacggcaacc tctgcctccc gagttcaagc aattctcctg cctcagcccc ccaagtagct 6366

ggaattatag acacctgcca ccatgcctgg ctaatttttg tatttttagt agagacaggt 6426

tttcaccatg ttggccaggc tggtctcgaa ctctggacct caggtgatcc tcccgccttg 6486

gcctcccaaa gtgctgtgag tacaggcgtg aaccactgca cctggcctat tttaattttt 6546

taaggaactc tgtatccagg gagtgcattt cgtctcagcc tttctctgta tgttcccgac 6606

tccctcttcc tcctcctgcc ttctccatgc tgtcctcttt gtccctgcag accattcttt 6666

cctttcccac acttctccct tccatgcctt tcctactccg gctcagctag aacacctgct 6726

tgctaggtag gtgggtgggg ctcaaaaatg gtgtggaaaa tatagtggtc tcagcctcag 6786

agatccatcc aatcagaggg atattatgta ggtgactgga gactataatg cttgaaaaaa 6846

ctagaatcct taaatgagga gggtggttga ctgggctttg agggaaagga cttctatagg 6906

ccagggaatg ggattctctc ctataagagt tagcactagc gctgcttctg ctgaaaccaa 6966

ggcatagaag cgggagatga tgtttgaatt catgagggaa gctggtaacg tggttggcag 7026

atattctaga tccctctgct tatcctgtct ctggccagca attcactgag atggcaggca 7086

gccgccagga gattggtggt gctcgcatct gtcctcaccc tattatgaac tggcttctgt 7146

gcctgctttg tctcctgtct tccacctccc tgcgagggag gccgttctac ctccattttg 7206

caggtgtgga cactgggtga ctcagcgata cgaaaatgct gagaggctaa cctgggtacc 7266

ttcccactgc acccctctac ctccgtggcc gggtaaggag ttgggctgtt ggtagagaaa 7326

agcttttgac cgtatcaacc agcagttttg ttcctggaca gctgccatgt gcaggtgatg 7386

tatggggcac agtaaatttc tgccagattt caattatgta gattcccttg actccatgga 7446

cagttacaag tcctggagtt tggggacctc acagttgtaa gtagccagca tacagaggtc 7506

agtcaagagg gacaagaagg tcacagggtg ggcctcttga gagaagagct gctgggtagc 7566

agatgatggt gcaggcagcc tggcctttgc tgtcatattg accggggttg gttccattgc 7626

aaccctgtct cttaatactt cagtagttag taaggctctt cctccaatgc cagttccctg 7686

ttggtaaaat agcaatgatt ttacctcttt cccagggtgc gtgttggggt ggctgtgata 7746

atggcaagta cctggcacat agtatgttct tgtacatgtt agctgcctta atcttggcaa 7806

gtgacatctc ctccttgaca attagctggc ctttaaaatg aaggccaaac tggaaaggtc 7866

aggcaggcca ctcctctctg gaggttcaaa tcccaggtgg tctaccatgt actaatcacc 7926

caaggtaaga gatttcttgc taagcctcag ttttcttatt tgcaaatcag aaattgatag 7986

tatttcattg ggtcaactac acagatgtgt gtaaagatat aaactatata taaaacctaa 8046

gctcttgcat tttggtgatg agtcacttga tttccatctc ttcacctatg aatggaaata 8106

tccatacctc tccaacctac tttcacaggg ctgtagtgag gtggtttatg gaaagggcaa 8166

tgaaactgta ctttaaatgc aatacaaatg caattgataa atgcaatata aatgcaactg 8226

ataagtgcat tataaatgca gtgcataaac ggttggagac catcgggctt gcatttttt 8286

ttttttttt tttagacaga gtcttgctct gtcgcccagg ctggagtgca gtggcgcgat 8346

ctcagctcac tgcaagctcc acctcctggg ttcacgccat tctcctgcct cagcctcccg 8406

agtagctggg actacaggca cccgccacca cgcccggcta attttttgta ttttttagta 8466

gagacggggt ttcacagtgt tagccaggat ggtctcgatc tcctgacctc atgatccgcc 8526

tacctcagcc tcccaaagtg ctgatgatta caggcgtgag ccatggtgcc cgaccaggct 8586

tacatttttt ttttttttc gtgagacgga gtctcgctgt gtcgcccagg ctggagttca 8646

gtggcacgat ctcggctcac tgcaagctcc atctcccggg ttcatgccat tctcctgcct 8706

cagcctcctg agtagctgga ctataggcac ttgccaccac gcccagctaa ttttttgcat 8766

ttttagtaga gatggggttt cacagtgtta gccaggatgg tctcgatctc ttgacctcat 8826

gatccgccca cctcggcctc ccaaagtgct gggattacag gtgtgagcca ccccgtctgg 8886

ctaggcttgc atttttttt ttttttttg agacggagtc ttgctctgtc acccaggctg 8946

gagtgcagtg gcgcgatctc ggctcactgc aagcccgcct cctggggttg atgccatttt 9006

cctgcctcag cctccccagc agctgggact acaggcgcac gccgccacgc tcggctaatt 9066

ttttgtattt ttagtagaga cggggtttca ccgtgttagc caggatggtc tcgatctcct 9126

gaccttgtga tccgcccgcc tcggcctccc aaagtgctgg gattgcaggc gtgagctacc 9186

acgcctggcc taggcttgca tttttaatgg attagttggg cagtgctctt tggatccgat 9246

gtgctgatgt ctaggactga ctgtttttgg tatgggagct tctgtctgca gtgtctagaa 9306

tgggggttgg agggactatg gggaacagag gtggccccag agcttggaca tctgggttct 9366

catgctggcc ttgctgctca taagcaggta acacagggtt gagttcctcc tttgcctggc 9426

cagccaacct ggaagaccta gccttctggc tgtcagtgtc cagagctctg gagaatccct 9486

gaggctacat ctctgggtgg gggaaggtcc tctctgtata gatggccctc caggggaagt 9546

ctgcaaggat ccagacttca gaataataga gttaagtatt aactaatcca agtggtccta 9606

gcacatttag cacatatggc atgctcaagg gtgtgcaatt gtgcagtcat caaagccagt 9666

gagggcaccc gatgccaccc atttcctgca ccaggaaaat ggacctttca cgtctctgaa 9726

cctgggatgc cagaatgcac tgaacttgat cctggcaaat aaattggaag ctgtttccaa 9786

aaaagtgata tgttcttgct acattgtctc cagagtttac tcctgagttg gggagcagca 9846

tctcattttg caaggggctg ttggattctc catagaaacg ctgggttggg ggagtgtctt 9906

ctgagtactg gggctctacc taatgtctga ttccatctgt ggtaccaaaa agccctctga 9966

atggcagcct gtctgtgcac aatgatgccg aagactttct tttccagcat ctaatgaggc 10026

atggccccag gaaagcagag aaggcggatg cttgagtctc tgagggtcag aatacgctta 10086

tggtgaaatc agaactgggg acaaggcaaa gtactgccat tgggaatccg gggtggatgg 10146

tcatgttaga ccaggggtgt ccaatctttt ggcttccctg ggccacttgg aagaagaatt 10206

gtcgtgggtc acacataaaa tacactaaca ctaaccatag ctgatgagct aaaaaaaaaa 10266

aaattgcaaa aaaatctcat aatgttttaa gaaagtttac gaaattgtgt tgggccacat 10326

tcaaagctct cctgggctgc aggttggaca agcttggcgt agactcttgg agggctttca 10386

tggtggaggc aggaaaggaa aagccccttt tacagatggg atagtaggtc cagagaaatg 10446

tgttggtttt cagttccaga gtaatttcac ctgctcccac cctccacact cacaaagagc 10506

aggctgggtg ggatgtcgat ttttttttct tctggagaat ggaccatctc cctatgcctc 10566

tcctcctcct tggctactaa cccctgactt ttccctcccc cttctctggc cactcctgtt 10626

cctgaacaaa gccagggaag atctgggctg gctcagcccc tgggtgtcag tcaagactca 10686

gagggagact ttgctcaaaa agccagccct cccggaaaga aatacagagc aactgtagtc 10746

aggaattttt ccctttcctt ttttctgtgg gcgggggggag ggacacttgg cagtctttgg 10806

ggcccttctg gctgccagcc ccacatctct tcctggctgt tgggaggcag ggcaggaggg 10866

aagattgggg agaaggggggc aggggctggc aaacacccag ctgctgggca gtggaattcg 10926

tccataggcc ctggtggagt tggcctctag tgaagcctgg ggaggtttct aattagtgga 10986

ggacggtgaa ggtgtggcag cctgtgttca ttcaaggaac tgagctctct gagccaggaa 11046

gtcaaaggga agaggggggac tgagtcccgg aaggccttgg taggtgggga ggaagttgta 11106

cctactcagt gtgttccatc ctattagaag ttgagaagag gggggtgttt cagctggagt 11166

gattcaggtt ggactctggg ggtgcagaga gtcagtcatt ggcactaagc aaagggagcc 11226

acctcaaaga ataagtgata taacagtgat aaacaaaatg gctttgttct agacattgcc 11286

tgcttctacc agaaagctca ttttgttttt tttgtttgtt ttatattctt cccctctccc 11346

acttcctgcc ccagctctgt tttagcctcg agattctctc cacttaacaa atgaagaaat 11406

caaggctaca ttcaggttca actgccaagg caagcacaga ccagattgta gtctgggtct 11466

ctagattctt ccagctctcc atttcaccag actcacacat gcagaaaagc agtgaattat 11526

atactgcttc ttggaagagc tccttttccc tgggaagtag taattaaggc atatcttctc 11586

aaaaatgttc ctttatgcaa ctctctagga agtttctttt aggtctattg tctttacttc 11646

tggatggagg tgccaagttg cctctgggtg gtgaacttgt tttctaaaat gcccgaggcc 11706

aagaacaaaa tgaaaatggc aattcccttg tgccctgtgc cctgccccgt gacttcctga 11766

taatgctgag cagttgctat agcaacagtg gtgggaggtg tgtcaggcag gataaagggg 11826

gtgggggtag gacaggttta ctggatcaaa gaggcttttg cggttgagta gggctgctga 11886

catgactgat agcctacttc tagggaaaga tgggggtccc atagagatgg ggaggggaag 11946

aaatgaagcc ctggggctaa aaggtcttgc ttgagagtct ttccttaggt gatgccaagt 12006

gaccaggtgc catctggttg gttaacaacc tgggcttggt tttggccact ggctacatga 12066

ctttgtgcag ttggcctcca gctctatagt tcatttatca agtgtgaggg ggataagaaa 12126

ttaccaacct catgtgcctg cttgataaca aaggtaaaca tgcaagccac gtagcacaga 12186

aactcagata atcaaatgag gtagtgcagg ttctgattct tgcaatacag gagaacagga 12246

gaatggcaga caccgctttt ggacgttttc tgattggtca tttcaagaat cctgttagca 12306

ggtgtgactt attcccttga gtgtggctta cctggacagc tgccactttg ctttggaaag 12366

cctccagtga gtccagagag gagtcaatgg ttgagagttt tccaggggggc agtcatgacc 12426

ctcgatgcca ttgtgcctgt ggttctacca acagcttaat gaccaaccat gtcccattta 12486

tgccccagaa gttaggactc ccacttgagg ctctgctggt tgtctctggt ttggttcact 12546

gcctcttaaa ccttgagagc atagccaagg aaggtaattg atgtgatata ctatcagaca 12606

aaattagtga actttagttg aagacagaga aaagcatggt gacctattta aagacactta 12666

tttatattca agacacttct tgtgctggtt tgagagtggc cggcaatcct ggacatcttt 12726

gcaggtttta aatttctttt tgcatcctgt gggttcaatt tgtcaggcct ggaatatctg 12786

agagcagagt tatggggatg gagaagagga cagtggaagg ggagggtggt gggaaaggtt 12846

ggtaggattc agcctttgtt acactatggg agtggggctg tgggattgca tcatcgaaac 12906

tgaattgttt gaatgtgggt cttaggacag gcatgtgtcc ttcccatccc atctcatggt 12966

tccacttgct cactgccttt cttgacttgc tcactgcctt tcttgactct taaccactct 13026

tgctgatacg ggctctagat gaggcaaagg ccagaagaat gagaaaaaga gttaatgtaa 13086

cagggtactc ttgggcaagt gctagatgct ctttgatagt gtattagtct gttttcacgc 13146

tgctgataaa gacataccca agactgggta atttacaaaa gaaagaggtt taatagactc 13206

acagttccat gtggctgggg aggcctcaca atcatggtgg gaggtgaagg cacgtctcac 13266

atggtggcag acaagagagc ttgtgcaggg aaactcccct ttataaaacc atcagatctc 13326

atgagacttg ttcactatca cgagaacagc atgggaaaga cctaccccga tgattcaatt 13386

acctcccact gggtccctcc tacaacatgt aggaattgtg ggagctacaa ttcatggtga 13446

gatttgggtg gggacacagc caaaccatat cagatggtca ccttttctaa tgtgtaaaca 13506

ttaattgagc acatgctgtg tgtcagctag gagaccatat gaaaaagcag ttgtcatacc 13566

acctgtaaca catgcttttt catgctggag ggggacaaga aaagcatttt agaaaggtga 13626

ataaagaaga aaccaaggca gcagagatgt atatacatcc tatgcaaata aacaggtagc 13686

tcttggtcac taaaagagct taaccaagtc tcttgggaag agcttgatgt taaataatga 13746

agtcccttca tggtaaataa agttatgaaa gtaagagaat caagttggac aggcttctac 13806

ttcttcagtc ttggaagaaa aaaaaaagca gtgtcttcct tcctaaaatg ttaattgagg 13866

acacagaatg tcttctcacc tgtccttcat tcactaagta tggaaagctg tacggaagct 13926

aagaccaagt ttatgtcttc aggaagcttc tgtccttagg aggactaaga aaatgcacag 13986

acctacagct ttttggcaat ggactcaggc tagagtgctg tgaagaaagt tgggtcaaat 14046

taccagaaga ctttacccag cagtgctttt ctttttaaa attttgtaaa atttttattt 14106

atttatttat tttatttatt ttgagacaga gtccccctgt tatgcccagg ctggagtgca 14166

gtggcgtgat ctcggcttac cgcaagctcc gcctcccggg tttgcaccat tctcctgcct 14226

cagcctcctg agtagctggg actacaggcg cccaccacca cgcctggcta attttttgta 14286

ttttttagta gagacggggt ttcaccatgt tagccaggat ggtctcgatt tgctgacctc 14346

gtgatctgcc cgccttggcc tcccaaagtg ctgggattac aggggtgagc cactgcgccc 14406

ggcccttatt tatttgtttt tttgagacaa gctctggctc tgtcacctag gctggaatgc 14466

aattgcacga tctcagctca ctgcaacctc cacctcccgg gctcaagcca tcctcccagc 14526

tatggcagtg cttttcatca atgcaatcaa gtgacctttg gacacttgga ggccatattg 14586

ttcatccttc ttccttctca aaaggtttct cctaagttat accagaggga gtctgtgcaa 14646

tttcttccca tctgcaattt atcagatttg ttttgttctt tcttatgaaa ttcttggatg 14706

gcttctaact tggatctctc ctgttttagt ttaaggctgt tttctctctg gtccttgggg 14766

agacagaacg gctccttccc ttcccttggg ttttcctggc ccttggacac ctccaccact 14826

cagcaccctc tggcagcacc aggatggatt tgtcccttgg ctttcagttc tcaggcctga 14886

gtcatcttga ttacttcatc ttccccttct taattctgtg accctcttct actacctttg 14946

ggtcttcttt ctgactcatt tccgagttct tcatggcata gagaacaggc caggagccca 15006

gggctaggtt ctgactggtg gacaagtgac cctggtaaca tcattgccct ggccagtgag 15066

agtggtggtg aagtccgcca tgtcactatc gtttggtcct cttccagtag ggcttcaggg 15126

acacttgagg ttactgatag ctctgcttga tgaggaagat gaggggtggt tggtgtgaga 15186

ggtgactttt ttgtttttgt ttttatttt gagacagagc ctctcactct gtcacccagg 15246

ctggagtcag tggcgcgatc ttggctcact gcaaccttcg cctcccgggt tcaagcgatt 15306

cttctgcctc agcctcctga gtagctggga ttacaggcat gcaccaccat gcctggctaa 15366

tttttgtatt tttactagag acggggtttc accatattgg ccaggctggt ctcgaactcc 15426

tgacctcgtg atccacctgt cttggcctcc caaagtgctg ggattacagg catgagccac 15486

cgcgcccagc cctgagaggt gacttttta aaagaatatt taaaaaatgc agaaaagtac 15546

aaagaacaat gtaataacac aagtctacca cccagaaccg tgggcggttg atattttgtc 15606

cactagcacg cagtacttac ttagtgttta tctttcttag ctaagcagac ttgcatatat 15666

gtgatagatg actaatatgc ctatgtttgc atatattgca tctatatgca tatatatgat 15726

agatgaccaa tattgtccaa attgttttgt gaaaataaat tgaaaatgca cataaaggtt 15786

ttattttccc tcttaacatt aggtgtctga gatctaatgt tgacatccat aggtcagaaa 15846

tattcttttt aactgctata tattgttcct ttgtattacc ctactggttt tatttatcca 15906

ttcacctgtt gaaacacttt tggatctggt tgctaatagt gtttttacaa ctacaagtaa 15966

tgctgtaatg accatttttg tacatgtgtc cctgtgcaca tgtagagaga agaattgctg 16026

ggtcaggcac attcttccca ttcccatagt ccctggttca agtttgaatc aaagcacagt 16086

ctactcactt gggcctcagc aactaggtca ccttgccagg ttcctggatg gctttagttg 16146

aacccaagga gaaatgtttg ggacccgtgg aaaggagtga aagctgatag tgctgtttgc 16206

catcttgata tcagtttctc tggctcttca aggacagctt ggctcacggg gccccagagg 16266

accttggctg aatttgagca tgtttttgtt ttctcaaact gtctgcctca gccctgaccc 16326

tccccaggtc tttgcaattc caggaatgac cctgatgttt ataaaggagc cctttggaga 16386

gtgcacaggc agggaggctc tggagaggct gtgctgtttc tgttcactcg gggggagccc 16446

gagatgagat ggcagcatga gggcatactc tagaaacctc catgacgcca ggccaggaga 16506

ggcccctctg catgtatcca agcaagcata tgtgtctgtt tacaaaccca tgagtggatg 16566

agcttacaag gctgtgtctc tgcaaagctt atacccgaga ccacctctct tgaaaagaag 16626

aggcattgtt ctcgcctcac caggaagttc ttcaggaaac agacagaaca gagagaatca 16686

caaaggcatg ggggcctggc aatgccagag cttgagagga gaggtatcat tcaacataca 16746

gattccttag tccttgcctc tgactttctg gggctggggt ccagggctct tatttttagc 16806

aagctcatca gatgatttct gatgaagagc tatgtttggg aacttctgac tgggtctaat 16866

tccctgaatt tacagatgag tacactgaca cctgcgcaag ttgaacaccc tgcttagcgt 16926

cactcagaat cccaggggtg gtggtgaagt caggatcgaa cctcagtcct tcctcccaca 16986

ctgggactgc ttggccaagg cttcgaaccc aaaagctggt gctccaaagg aaatccaaaa 17046

agagtttgca gagtggtttg acgtggaggg gattcatcgg ccaggtgggg gctctctgcc 17106

ctcccttggt ggtgcatcat ctcctgtagg tgttttgtca ccgtggaagg gcagtagcct 17166

gtgcctgttg ggcttgccac tcaacagagg tcactgagct cttccgcact cagctgggtg 17226

attggcctgt ggtagctgca ggtggaggga gatgattctc agaagagaaa ggccagttct 17286

tgcagtcgtg tgatcttata acctagggga cttgtggatc tgaaacatct gagaaggctg 17346

tttgcagtta tttactacca caaaggaaac tgggagagga gggaccacct gttggaaaac 17406

catgctcaga agtgactggt ctgttgctga aggatggaag gatatcaggt gttgcatggc 17466

tggctccctt gcaggaagga agggctcagg gggtgagaag gaataggaag ggagtttaca 17526

atgaatgaag ctgggtgtcc tcaggagata gatggaccag ttggggtgat gtcattaact 17586

aagctggaat gtgtcctgag tgccagccta taccgcgaca gacatctgtg cacactattc 17646

agaattccag tgtgagaggc cactgccccc tctgaccttt tgaaaaaaaa agtcagggga 17706

agacacataa acagtgtcac tcacccacaa agtctcagga cctcaggtta aactttagct 17766

aaagaagtaa caaatagaac ccccatttta ccttcctctt tcctaacttc tgcctccctc 17826

cctcccacca gcatggcctg aattactaat cccagcctgt tggtgggctg tgctattgaa 17886

catagtggcc acatgctaga tggagtttgg tgaccaccat cagccatcca aagaggactg 17946

gcttcaaatt ctctccaaca aaagctgttt cttgtatgtg ctggcgtgtt tatgagccct 18006

gtctcaactg agcatgcgtg gtgattatga gctactgcct gacctccttc cctttgtccc 18066

tcccgctaaa ttctgtaaac aaagttaaag taatttaggc aaattcattc ttgcttaagg 18126

acttctatta aatggcagag agctttgata tttagcaatg tataaaattt gtacaatcaa 18186

aaccaagctt tagaagaaat atgcaataga tcatattgca gaactctcat gactcatttt 18246

ttccatgcat ttggccattc tacaaatcag acattaccta gccttccccc cacatggcag 18306

atataacttt tatttaccca ttcaagaaat tgattgttca cacattcctt tttctttgaa 18366

acaaggtctt gctttgtcgc ccaggctgga gtgtagtggt gtaatcttgg ctcactgcaa 18426

gaagtgaact ccgcctgctg ggttcaagcg attcttgtgc ctcagcctct caagtagctg 18486

ggacctcagg catgtgccac catgcccagc taatttttgt atttttggta gagacaggtt 18546

tttgctatgt tggccaggct ggtcttgaac tcctggcctc aaatgacccg cccacctcag 18606

cctcccaaag tgctgggatt acaggcatga gccaccatgc ctggccacca cacatttttt 18666

tctaaagctt ggttttggcc acagtgagag tttcttgggc tgtcaggggt aaaaatacca 18726

tgtcttgttt aacagcagcc atcattttag cagaattttt ttttttttt tttttgagac 18786

tgagctttgc tcttgttgcc caggctggag cgcaatggcg cgatcttggc tcactgcaac 18846

ctccgcttcc cgggttcaag tgattctcct gcctcagcct cctgagtagc tgagattaca 18906

ggcatgcgcc accatgcctg gctaattttt atgtattttt agtagagacg gtgttactcc 18966

atgttggtca gactggtctc gaactcctga cctcaggtga tctgcccgcc tcggcctccc 19026

aaaatgctgg gattacaggc gtgagccacc gcgccccgcc ccttagcaga attgttaata 19086

tcatttttca gacttgagaa attagagctc agagaggtta aaggttctga ctcgtgagtg 19146

aatctcaaag ggcaaagctg gcattcaggg ttcttgactg ctgggttcat tgctcattcc 19206

agtttgttaa tgaatcgtcc actgctatgg atattaatgg agtcgtctag agagcagggg 19266

gaggaggatg gtggttagta tttattctgt tgcccaacat atgtcaaggc taagtgcctt 19326

ttgggagaaa ttaacaaaga cggagccatg tgagttttct gcccagctag ggtggaacca 19386

ggaagctgcg gccagccttc attcccctgg cccattactc agaattaact cttgaaaggc 19446

ttggaatgct attacctgtt gcaagaggtg agttcattgg aggaatgata agttcagggt 19506

tggcttagat catctctgtg gtactcctct ttcccaaccc tgctcttaga ttttaagata 19566

atcatgagat aggaggagaa ggctggaaaa tggtaagtct ttcttcagaa taaagaaaag 19626

aagaaagaag agagacacat gagctgttag gaagatgtga ttgttctagg tagggaaaag 19686

gtcttatttc ctcaccaggt ctggataagg accggggcca cacactctga ggaggaagga 19746

ccagggaggt gcttttagat tcttggggct tagtttccta agaaagcttt ttaaggtctc 19806

tcttagatca ggaatcacca atttccaaat agagcactga ggtcagaaca ttgtgctagg 19866

agaaaatgag attttgtggg agtccgtttc tatcaaacac aagagtttga tttgttgagt 19926

aagatcattg atgcaacaaa caggattaca ttggaaaaat gagctgtgcc aaaattaggt 19986

ctgaaaacct gggtatgaat gcatagaagt gaacaataaa ggctctttct cttgggctgt 20046

cagcttccag ctgctctcct tggctaaatt agtaagggtg ggccaaaaga caatgtaggg 20106

accccatgat aattgtgaag aatggcagag aggttgtttg ctcctgtgaa ttcaaatggt 20166

agtccatctg ggctcaggtt tgccatctgt aaaatgggac aacgatccat ttcttttcat 20226

cccatctgcc ttcctcatct ccagctctcc tctaggggtc ttaggtagaa aggtgtggca 20286

gtctacccaa ggcaggggtg ggagttctcc cctgcccttc aagcccccga ctaatggctc 20346

tggggttgtc tgtcatctcc ctgacttctc tttggaaaag cttgaccttc tatatgtgaa 20406

atcccttgct tctgtgttga cccatttccc taggtgctat gtagggaaga agagctaatg 20466

cttactaatc tacagggaaa gtctcttcag gggccttgga ggtttctatc aaccatttct 20526

aagcccatga gccttttcca ttgatgtcat ttggggttgc taagaaccat ttcttatagt 20586

gacctttgca ccatagccca tggagttgca tacataagtg gttgcttaat aaatattatt 20646

ctgaacttaa ctctcttagg cagaaaaaaa atacaacttg gtatttttgt tctgagggca 20706

aattccatct ctagctggag ttcgacatct ggaatgtttt cctgtggatt gttcttgcag 20766

cctataagag gatatatggg aagaacttgg taacatgtga taatgccctt gtctgactga 20826

ggggaggagg ggcgctatgt gggtacttcc tggtaggcag gaaatggaag tggctaagag 20886

ctaaccccat gtgtgcctac tgtgtgccag gccctctctg tagatgcagg cagcatctct 20946

ttggatcctc ccagttgccc ataggaggtg catactgttt gtccttttac agatgaacaa 21006

tggaagcaca gagagcttaa gaaagttgtc tggtgctgca cagctcagta gtaaaagctt 21066

gaattccaac tcaggcagtt agattttggg gtttgaatct ctttggctgg gaacagtaac 21126

aaggaatgca ggcttttctc agcctacaat gctttctgat ctctgcctgg tcaactcttt 21186

catttccctt ccctgggctt ctacttcaag atgtggagtt cagatctggc tacgggattt 21246

gagtccagaa gcctcttcct gagaaagaaa cagatgctga tgcatggtct tgttaaacca 21306

tctttaaaag accttgcatc agcaaaccag gcctgtggat ttcttcatct tattggcatg 21366

ttcgttggca ctgctgagat caataaagaa atggaatgaa acataccctt gaatccaggt 21426

ggcctctgtg agggtggaca aacttgacct tcacctcaga gagtttggga catctcaggg 21486

caggaaatca actttgtcct cctggaggaa cactgccaag ttgaccccag cctcctggga 21546

cctggagcca gggttgcata tgacatctgt gggccctgga cactgcctct gtgggccctt 21606

ctataaaaac atatattttt aaaatttaaa aattattatt tttcgttttt gagacagggg 21666

ctcactctgt cacccaggct ggagtgcagt ggcatgatca tggctcactg cagcctctaa 21726

cccctgggct caagtaattt cccacctcag cctcccaggt agctgggacc acaggtgtgc 21786

atcaccatgc ctggctattt ttttatttt agcagagaca aggtgatata gttaggcttt 21846

gtgtcgccac ccaaatgtca tcttgaattg taatccccat aatccccagg tgttgatgga 21906

gagatcaggt ggaggtaact gaatcatggg ggtgggtccc ccatgttgtt ctcgtgatag 21966

cgagtgagtt ctcctgagat ctgatggttt tataaggagc tcttccccgc ttcacctggc 22026

acttcttcct gccaccttgt gaagaaggtg ccttgcttcc cggacctttg gcttctacca 22086

cgattgtgag tttcctgagg cctccccgcc atgctgaact gtgagtcaat taaagctctt 22146

tcctttataa attacccagt ctcaggcagt tttttatagc agtgtgaaaa tggactaata 22206

cagtcaggcg tggtggctca tgcctataat cccagcactt tgggagaccg aggtgggcct 22266

cctgtggcaa ggagttcaag accagcctgg ccaacatggt gaaatcctgt ctgtactaaa 22326

aatacaaata ttaaccaggc gtggtgacac atgcctgtag ttccagctac tcgagaagct 22386

gaggcaggag aattgctcaa acccaggagg cggaggttgc agtgagccaa gatcgtgcca 22446

ctgtactcca gcctgggtga cagagcgata ctctgtctca aaaaaacaaa aaaaaaaaat 22506

caaaaaacca tattttataa cttgtttaat aaatgcaaat gtattaatat tatatagtaa 22566

gacttttaa cctaaaagtt attttttttt tctgatgtta atagaaataa aaatatgtcc 22626

taggcactgt gcctccagag cccaatggac aagttggccc tgcctggcgc cttcaggccc 22686

gaatgcacat ttatctgtgt gtgcccctct ccttttttcaa atgtgtttcc agagattgca 22746

gagggccttc ttctgaggat taatgttgga ttgacctctg ggacgcagcc aagggatggt 22806

tctttggcca ctgtctgtgg tacaagaggg cttcgtgaag atggggagct gcaggctgaa 22866

agtcctctag gaagcccagg gaagatcctg tgccctgccc ccaagaaact gtggactatt 22926

gtggctgcag atggaagggt agcggagggg tcttagccgt ggatgcggaa tatagtggag 22986

ctagatttgc tttatatggg gcctggagcc atttcagcaa tgggctatca gcttacaggt 23046

cttgggctaa aagcctgact ccagaagagt gattctcacc ctgggcagtt ttgcctctcc 23106

ggggccattt ggcaacgaat gaagacacat tttttgttgt tccaacctgg ggagggtgct 23166

actggtttct ggtgggtaga ggccaggaat gctgcccaac atctggagta cactggatgt 23226

tccaggcatt gagttcgggt gagcggctcc ttcccagaag ctcatggata gaggtacaca 23286

tgagctgcct ttcccgtctc caaccagaca acgtgttgtt atagccataa tagagactgt 23346

ctgccttcac ccaagccact tcccaagagg gcagagccag gtgagcctct gagcatcaca 23406

gacttcatgt ccacagattg gagagttcca gaaatggtgg cagggtgcag taccagcaga 23466

ctcccagagc tcctgtcttc tctaacaccc tgtgagcaga atgcggtgtc gaggtctcca 23526

gaaacccttc ctgcctctgt gggcacaaag ctgggtaggg tgcggtaagc agggtgatgc 23586

caagggtgat ctcaggtacc aagggttcgg gtgagggcga actggaggtt cccacttcat 23646

tttgtctttc tcattttgat cctgagtagc tgggagggag tcaaagggct cttctctatg 23706

ttccaatgca ggtggtggga gtggggagac tttctagcta ctttgctctc ctcctgtgtt 23766

tttggtgtct gcagagggca ggggcaggga gcttgagatg gtcccactca ggtctgggga 23826

cagagaggcc tgtctgggaa gggcagcggg gcggaggcag ggaggctggg gatgagtgcg 23886

tgcgctgcac ggcatctgtg tgctgcgtta gaggtcagtg ttgagatgga aagttgactg 23946

aacctgatct cttctgtttg ttgtccaggt ttgtctgatc tcagactttc cccggccttt 24006

ctctttcacc cttttgtttc tgacctctgg ggtccatctg aagagggaag gggggctctt 24066

ctaggagggt gtggacctgg tggcgctgaa gaaatctgcc gcacggtggg gtgactcagc 24126

gtccctgttc ttgggaacct ggcaggcagc tgagcagggg cactcactca gcctgcctct 24186

ctgggtccct gggtgtctgc tgctctccat gtggaggagc acagagggca gggcaggatg 24246

gatgaggaaa gctgtgggct ggtgcctggg tctgattaga catggtggag caggactcca 24306

acctagcctg gaggcgcacg ctaaatcggc cttttcctgc cggtggtgct gggctgccgg 24366

gctgctggaa caatttaatc tgtgctggct tcatttatgg tacaacccag gctccccaga 24426

gcccctgtgg gaataaccat tctaaataaa atatttttcc aggtttcagg tttatacatt 24486

taacatttaa attttttttt tcctggagtt caacaatttt ggataaaagt cttttaaaat 24546

gatgaatagg tagaatttta attccctccc ctttttgatg tcccaagatc atcgtcatca 24606

acgtaattct tgttttgggc tgtaagacat ggaaatcccc agggttcatt gaggggcatt 24666

cattgtttgc cacacacaaa atgtacccgg agtttgtatg ttttgttttc ttagtgtcct 24726

ctttcctctc tgccatttga tccttcttgc tactgtccac atgcattcct tctttgtccc 24786

ttcaattttg ttgactttct tctgcggtgg ccttggaaac cagacatcca gaagccaggc 24846

ttcctggagc agcactgttg tgaactgtct cgaggctcct cacgagatga ttctcccctc 24906

atcttgagtg gaaagtgcca acatttgaac atgctcagcc gtggtaacag ctattaggat 24966

gttcagagct tattttcctt ggtttccctg ttctcggttg ggttccatgc tgctgaagta 25026

gtagataact ggggcattcg tcattgtcat catcatcagt gtaataacca gcagcaacat 25086

ggatatagaa tttctgtgct ccaggtcgtc ttttatttat ttatttattt atttatttat 25146

ttatttattt ttgaggcagt gtcttgctct gtcgcccagg ctggagtaca atggtgcgat 25206

ctcagctcac tgcaacctcc acctcctggg ttcaagcaat tgtcctgtgt caccttctg 25266

agtagctggg attacaggtg tgtgccacca cacccggcta attttttttgt atttttagca 25326

gagatggggt ttcaccatgt tggccaggct ggtcttgaat tcctgacctc aagtgatcca 25386

catgcctcag cctcccaaag tgctgggatt ataggcatga gccaccacac ccggcctaag 25446

tcatctttta actgcttttc tttcgtcata ctctcaacag ctctatgaag aaagaactgt 25506

tattatccag caacttacag ttaagaaagt agaggcacag agagggtctg tgtggggaga 25566

acacgtgcag ctgatataga aaatggcaga gctggacttt aacccggcta cacccttta 25626

accagagctg ctaccctggg gacggcgtgc tggcttctaa aggaagctgc tgcatggcaa 25686

gtgaagagtg actagttctc aggcttggag ggaggggaaa ggacagggac agacccccat 25746

caaatgttcc gtatgctggg gcctatgctg ggcacgcctt gacagcaggt gctgctggag 25806

tgccagcacc cggtgctgtc cctggcatgg aagaggtgcc ccctccctta gttggtgagt 25866

gagtgaatac tcttgtatag agttactcta agaattattc atttgtgttc taagggctgt 25926

aagtttttga aggtgcatta atggggaaga aaaaatgtgc aagattggtt gaaaaacacc 25986

cctctgcacc cacccagatc ttgttgggac ctgcaggaga gcagctctgg gtgggagggt 26046

ctggggctgt ttgccgtgtc tccaaggggt tggggctggg gctcggcggt gtgaggccct 26106

gcccttctgg agcttgctgg ggcttgtctc ctggttacag ctctgccgtg cgattcagat 26166

cctctctgcc tcctttggct ggacgtgcct cagccagagt gagactgtgg caggctggag 26226

ctgttttcaa ggcctccttc acctaggctg tgcatacaaa aggttgaaga aagtataagg 26286

agacttaatg gacgcttgtt ggacagacat cacgaagctt catgccctgc agtaagctga 26346

agctatgaag tcgaaggagt catagttcct gttcttgggg ctctcagtat aatggggttg 26406

gggtgggctg cacagacaga cttgggaacg agggaatgtc agtgtggcag agctgtagct 26466

```
gcctcggagc agcactatgg gaccttaggc agggccaggt gggcaggtac acagtgacag  26526

caggggggatg ctgggaccag gttctgggtg gcgacgtggc cttccacctt ctcactcctt  26586

tgccatatat gctgcagccc ggtgatgtct gtggcacgag ttttaaaaaa ctcaactaat  26646

tggtgctctt ccagaccaca aatgaagaac tttaaactgg gggttatagt gggagaattt  26706

ggggggggtcc cagaaaggac ttaaacccat tccatggtca gggcctttca aatgcattca  26766

tccctctgtg tgcctctgtt ctgccccagc cttgtaagag ggtccctgtc attcttctag  26826

agccagacag cctgatttca aatcctgaat ccatcgcttt tcagatttgt gatcttggtc  26886

aagtaactga cctctgtgag cctcaaattc tccatctctg aaagttgggg agggtgctag  26946

tagtacccac ctaagagtta gggtgaagct gaaattgatc atccatataa attgctcagc  27006

acgggccctg acaggccatc agcactcatt agccgcgcgt catctccacc ctcatcatca  27066

ttacggtcca ggtggctccc ggagggcacg tctctctggg aaactggacc tggagctttt  27126

cttcccttcc tctcttggca g ac  aat ggc gat gac cac tca gaa gga ggc      27176
                        Asp Asn Gly Asp Asp His Ser Glu Gly Gly
                                150                 155

ctg gtg gag aac cac gtg gac agc acc atg aac atg ttg ggc ggg gga      27224
Leu Val Glu Asn His Val Asp Ser Thr Met Asn Met Leu Gly Gly Gly
        160                 165                 170

ggc agt gct ggc cgg aag ccc ctc aag tcg ggt atg aag gag ctg gcc      27272
Gly Ser Ala Gly Arg Lys Pro Leu Lys Ser Gly Met Lys Glu Leu Ala
    175                 180                 185

gtg ttc cgg gag aag gtc act gag cag cac cgg cag atg ggc aag ggt      27320
Val Phe Arg Glu Lys Val Thr Glu Gln His Arg Gln Met Gly Lys Gly
190                 195                 200                 205

ggc aag cat cac ctt ggc ctg gag gag ccc aag aag ctg cga cca ccc      27368
Gly Lys His His Leu Gly Leu Glu Glu Pro Lys Lys Leu Arg Pro Pro
                210                 215                 220

cct gcc agg gtcagtgagg gtcaggtctg gtggaagggg tgggaggaca              27417
Pro Ala Arg

aggaaagtgg ggtctcagct ggaggagggc atcctaataa gaccctcatc tggtgtgacc  27477

tgtaggcaaa gcacttttct ttccacaaac atagttgtgg aacatagtta gtgacctttg  27537

gggcagtcag tatctctggc tgtgaaatag actggacatg tggtttcttg atgttctttc  27597

cttctaaaac cttcacattt ctgtatggtt ttggaacata gacacagaaa atgaatccca  27657

gagatgaaga cagagtccct gaaacttgct tactcttttt tttttcctcc ccctgagatg  27717

gggtctcact ctgttgccca ggctggagtg taatggtgcc atcacacctc actgctgcct  27777

ccaactcctg ggctcaagtg atcctccttc ctcagcctcc tgaacagctg ggactatagg  27837

catgcaccac cacacctggc cagttttgta ttttttggtg gaaatgtggg tctcactatg  27897
```

```
ttgtccaggc tggtctcaaa ctcctggact caagcaatcc tcccacctcg gcctcccgaa  27957

gcgctgggat tacaggcgcc agccactgcg cctggcctga aactcagtct tgttgccata  28017

tgcttctctc actgagctac ccctcccctc tttcctgccc agctctgcca tcttgggtcc  28077

tggccctggt tggttttatg gaaagaagtg gagcttgggc atggggtggg ggcatggaga  28137

ctcgagacgg ccacttagat ggatatttgg aggcagagcc cattctgaca catgaaagcc  28197

acggtcacag aacaaggaga ggagtgaaca ttcccatcca tacaattttc tttaggggct  28257

ggggttgggc acccagcatg gcagggagat ccccgggcag gctgtcagaa gtccctgtgc  28317

cccctgctgg ctgtgtgacc atgggaaagt cccctccctg tgcagcagct tctcgctgag  28377

cttgcgtctg gcgcgtgctt cgtgggcctg ctgtcctcac tccgggcgtc ccctgctgtc  28437

tgtgcgtgca g act ccc tgc caa cag gaa ctg gac cag gtc ctg gag cgg   28487
           Thr Pro Cys Gln Gln Glu Leu Asp Gln Val Leu Glu Arg
           225                 230                 235

atc tcc acc atg cgc ctt ccg gat gag cgg ggc cct ctg gag cac ctc    28535
Ile Ser Thr Met Arg Leu Pro Asp Glu Arg Gly Pro Leu Glu His Leu
       240                 245                 250

tac tcc ctg cac atc ccc aac tgt gac aag cat ggc ctg tac aac ctc    28583
Tyr Ser Leu His Ile Pro Asn Cys Asp Lys His Gly Leu Tyr Asn Leu
   255                 260                 265

aaa cag gtgagcatgg tgccagcctg.ggccagagca gctcctcctc agccctgggc     28639
Lys Gln
270

cccagatgtg ccttgtttct gccccacccg cctatgatcc tctgaggtct gagctgagtg  28699

agagactcag actcctgtca ccatggggat tgggatgcca gctgccaggc cggggagggt  28759

gcagctcttc tccctgcctc cgacgggtca gttgctggct ccacttgccc agctgcccac  28819

cggcctcact tggagtccag gttgggagga ggctcttgga gaagggaagt ggcttaaacc  28879

aagcatcttt gctttgatcg gacattgaca tcacctggag ctcttcctgg gcccgccccc  28939

aaattgccat gtggtaggtc aggtttacag ctctaacaag cctcaggtga ggctcataat  28999

ctgtgagcca cactcccacc atcattgatt tcaaccactg gccgagttgt cttcatggtt  29059

tcgaagtatc tcagggaatg aagagtcagg catggagaaa tgccgttgga aagattctga  29119

aaagccatct agaacaccat tgtaaataca ttggagaaaa tgaagtttaa gtaaaccact  29179

caggagccct tctccctcct ctatccacct tctgttcata tgcattttaa cagtcactgt  29239

gcatatgcaa tattttccct gtaacatact ttttttttct ttttttttga gacggagtct  29299

cgctcttgtc tcccaggctg gagtgcagtg gcgtgatctc agctcactgc aacctccgcc  29359

tgccgtgttc aagcaattct cctgcctcag cctcctgagt agctgggatt acaggtgcct  29419

gccaccatgc ctggctaatt tttgtatttt tagtagagat ggggtttacc atgttggcca  29479
```

ggctggtctt gaactcctaa cctcagatga tccccccacc tcgacctccc aaagtgctgg 29539

gattacaggc atgagccacc acacctggct gggcattggg atttttataa agctaccctg 29599

gtggttctaa tggcagtctg ggctgaccta aagtgtggat gtgtgaagcc aggagatggg 29659

aacagggctc tgctgatggc ttaggggtct cacctctcca ccccccttcac ccaggcacct 29719

tgaccctcta aatattagtt tagtattaaa agctatttgt gtacatgccc tctaaatgta 29779

agctggctaa ctccatttgg agctcagaag aaaaagtctc gatcacttct catgggaaac 29839

aggctgtctg ctgcaccttt gggtaactaa agtgatttgc aggagtagtt agcttatgca 29899

taaaacagag accttggagt gggaatgcct ggcttccatc tcagctgtgc ttcttaccag 29959

ctgtgtgacc ttgggcaagt tgcttagccc atctgtgctt cagttttatc atctgtaaac 30019

cagggataat aatggcatct aattcacaaa gtggttgtga ggattaaatg actatatttg 30079

ttagaaactg aactagcagt gtgtcttgca cacacagtgc ctggcacaca ctgattttga 30139

ttgcaacatg aggttttaca cctgacgcta tgactaagct tcagtatgct gggtcccccc 30199

cactggatgg gacctctttg aggctgggga ctgagtcttt tcatcttcgg gtgtcctgaa 30259

tggccttgct aatcagaggg gtccctgggc gggcagcatc agcatcacgt tggagctcgc 30319

tagcgatgca gaatttccag ctctccaggg acctactgca tgggaatctg cattttatca 30379

aggtttcctg gcgcatattt gaagttgcga agctccaccc ggcagcgcat gggtagctgc 30439

aaggggtggc tgctcagtgg acgccgggtt gagggacaga aggaaagttg ctggctgcgg 30499

gctcctccat gctcttctcc tctctcccca g tgc aag atg tct ctg aac ggg 30551
Cys Lys Met Ser Leu Asn Gly
275

cag cgt ggg gag tgc tgg tgt gtg aac ccc aac acc ggg aag ctg atc 30599
Gln Arg Gly Glu Cys Trp Cys Val Asn Pro Asn Thr Gly Lys Leu Ile
280 285 290

cag gga gcc ccc acc atc cgg ggg gac ccc gag tgt cat ctc ttc tac 30647
Gln Gly Ala Pro Thr Ile Arg Gly Asp Pro Glu Cys His Leu Phe Tyr
295 300 305 310

aat gag cag cag gag gct cgc ggg gtg cac acc cag cgg atg cag 30692
Asn Glu Gln Gln Glu Ala Arg Gly Val His Thr Gln Arg Met Gln
315 320 325

tagaccgcag ccagccggtg cctggcgccc ctgccccccg cccctctcca aacaccggca 30752

gaaaacggag agtgcttggg tggtgggtgc tggaggattt tccagttctg acacacgtat 30812

ttatatttgg aaagagacca gcaccgagct cggcacctcc ccggcctctc tcttcccagc 30872

tgcagatgcc acacctgctc cttcttgctt tccccggggg aggaaggggg ttgtggtcgg 30932

ggagctgggg tacaggtttg gggaggggga agagaaattt ttatttttga acccctgtgt 30992

ccctttttgca taagattaaa ggaaggaaaa gt 31024

<210> 10
<211> 325
<212> PRT
<213> Homo sapiens

<400> 10

Met Leu Pro Arg Val Gly Cys Pro Ala Leu Pro Leu Pro Pro Pro Pro
1 5 10 15

Leu Leu Pro Leu Leu Leu Leu Leu Leu Gly Ala Ser Gly Gly Gly Gly
20 25 30

Gly Ala Arg Ala Glu Val Leu Phe Arg Cys Pro Pro Cys Thr Pro Glu
35 40 45

Arg Leu Ala Ala Cys Gly Pro Pro Pro Val Ala Pro Pro Ala Ala Val
50 55 60

Ala Ala Val Ala Gly Gly Ala Arg Met Pro Cys Ala Glu Leu Val Arg
65 70 75 80

Glu Pro Gly Cys Gly Cys Cys Ser Val Cys Ala Arg Leu Glu Gly Glu
85 90 95

Ala Cys Gly Val Tyr Thr Pro Arg Cys Gly Gln Gly Leu Arg Cys Tyr
100 105 110

Pro His Pro Gly Ser Glu Leu Pro Leu Gln Ala Leu Val Met Gly Glu
115 120 125

Gly Thr Cys Glu Lys Arg Arg Asp Ala Glu Tyr Gly Ala Ser Pro Glu
130 135 140

Gln Val Ala Asp Asn Gly Asp Asp His Ser Glu Gly Gly Leu Val Glu
145 150 155 160

Asn His Val Asp Ser Thr Met Asn Met Leu Gly Gly Gly Gly Ser Ala
165 170 175

Gly Arg Lys Pro Leu Lys Ser Gly Met Lys Glu Leu Ala Val Phe Arg
180 185 190

Glu Lys Val Thr Glu Gln His Arg Gln Met Gly Lys Gly Gly Lys His
195 200 205

His Leu Gly Leu Glu Glu Pro Lys Lys Leu Arg Pro Pro Pro Ala Arg

210 215 220

Thr Pro Cys Gln Gln Glu Leu Asp Gln Val Leu Glu Arg Ile Ser Thr
225 230 235 240

Met Arg Leu Pro Asp Glu Arg Gly Pro Leu Glu His Leu Tyr Ser Leu
245 250 255

His Ile Pro Asn Cys Asp Lys His Gly Leu Tyr Asn Leu Lys Gln Cys
260 265 270

Lys Met Ser Leu Asn Gly Gln Arg Gly Glu Cys Trp Cys Val Asn Pro
275 280 285

Asn Thr Gly Lys Leu Ile Gln Gly Ala Pro Thr Ile Arg Gly Asp Pro
290 295 300

Glu Cys His Leu Phe Tyr Asn Glu Gln Gln Glu Ala Arg Gly Val His
305 310 315 320

Thr Gln Arg Met Gln
325

<210> 11
<211> 10884
<212> DNA
<213> Homo sapiens

<220>
<221> promoter
<222> (102)..(1905)
<223>

<220>
<221> mRNA
<222> (1906)..(2440)
<223> Exon 1

<220>
<221> mRNA
<222> (5726)..(5952)
<223> Exon 2

<220>
<221> mRNA
<222> (6497)..(6616)
<223> Exon 3

<220>
<221> mRNA

<222> (8212)..(10775)
<223> Exon 3

<220>
<221> CDS
<222> (2038)..(2440)
<223>

<220>
<221> CDS
<222> (5726)..(5952)
<223>

<220>
<221> CDS
<222> (6497)..(6616)
<223>

<220>
<221> CDS
<222> (8212)..(8334)
<223>

<400> 11
ctgcagacct gggacctcaa gaattgcatt tgatgccgaa cccagctcta atttcagagt 60

caaggtctct gcgagtattt aaggaacgga tgtaaacctg ggggattcgt tttgtttcct 120

tcaattttcc aatgaaatca gagatcctgt tcttgggtgt caacgcagat actagaagga 180

ggtgatacaa gagaaaggaa acagcaagcg acgattatgg cacggtttcc tgtaaacaag 240

gttgagtgta gccacagcct gagcactgtg ggagaagagc tcataagaaa atgacggtgc 300

tgggccttcg tcaccccggg gccctccatt gttcttgtct ttggtctctt tttatttgta 360

gaggtccaat tatttattta tttagtacaa gagggaacga aattgatctt tccattctaa 420

aaggagagta tatatgtata aaaggaagct gtatagatat gggggaagag gtggacaggg 480

ggaaaagggg agaggacgag agagagaaag ggagggagag ggacaaggag agacactggg 540

cgagagatcg attaggagag acagaaatga tgaatgaaga ttaacttcac ccaaggcttc 600

gtcgctggag gggaatggag gagctcctga tttgctatta ctactccaaa ctgcaaaggg 660

ctccttcaag tcacctatcc acctcctaag gcaagcgtcc aatttcaaca gcgttcagga 720

aagtctcctc ccgcggaggt ctcaccgctt cccactccac ccccacaaac tctttggaaa 780

agtgccttga aaaatttaat cctcaatcca atcctggacc accagcgtcc tctgttggtc 840

accgaaggag ggggtgcgca gacaaaactg aagaaactcg agtgccagag aaggccgaca 900

ggagttacag cgacctcagc gcgcaattgc gccccgaact ttactgaaaa gtgtttagat 960

tgcagagata agctagaatc ccaacgcatc gagaatacag taatacgaag tcgccttcaa 1020

aaaatgacaa tgaaaattgc ctattaaagg actatttggt taattacgtt tcagcagtgc 1080

ccagtttatt gtctttatta ttcttttgtc gtgggtgtaa actccatttg aaaacataat 1140

cagggagaat acccaagaca agaagaacag ttgtcattta aaatatttga aaagccctgc 1200

cttaaggagc attcgcttgc cggtccactc ttaattgggg acttgcggtg tagcaacacg 1260

tgagagtctt cttgcgttga gaagtaagcc tggaaaggcg aaggccccgg ggcatcttca 1320

gatgcgtatt tgtgggcccc tggggatata aacagcccag cgggtgtaaa ttaaaccccg 1380

cagtgccttg gctccctgag acccaaatgt aagtcagaaa tgtcccaaga cttcgcctgc 1440

caacggaatt aaattttaga aagctccacg aggtacacac gaatgcggag cgctgtatgc 1500

cagtttcccc gacaccggct cgccgcaggg agacctcacc ccgagagcgg aaggggtaag 1560

ggcggcgggg tcaaggagat cgggggtgct gagttggcca ggagtgactg gggtgaccgg 1620

gggtgctgag gtggcctgga gtgccggggt ggccgggcac accttggttc ttgtagacga 1680

caaggtgacg ggctccgggc gtgagcacga ggagcaggtg cccgggcgag tctcgagctg 1740

cacgcccccg agctcggccc cggctgctca gggcgaagca cgggccccgc agccgtgcct 1800

gcgccgaccc gccccccctcc caacccccac tcctgggcgc gcgttccggg gcgtgtcctg 1860

ggccaccccg gcttctatat acgggccggc gcgcccgggc cgcccagatg cgagcactgc 1920

ggctgggcgc tgaggatcag ccgcttcctg cctggattcc acagcttcgc gccgtgtact 1980

gtcgccccat ccctgcgcgc ccagcctgcc aagcagcgtg ccccggttgc aggcgtc 2037

atg cag cgg gcg cga ccc acg ctc tgg gcc gct gcg ctg act ctg ctg 2085
Met Gln Arg Ala Arg Pro Thr Leu Trp Ala Ala Ala Leu Thr Leu Leu
1 5 10 15

gtg ctg ctc cgc ggg ccg ccg gtg gcg cgg gct ggc gcg agc tcg ggg 2133
Val Leu Leu Arg Gly Pro Pro Val Ala Arg Ala Gly Ala Ser Ser Gly
20 25 30

ggc ttg ggt ccc gtg gtg cgc tgc gag ccg tgc gac gcg cgt gca ctg 2181
Gly Leu Gly Pro Val Val Arg Cys Glu Pro Cys Asp Ala Arg Ala Leu
35 40 45

gcc cag tgc gcg cct ccg ccc gcc gtg tgc gcg gag ctg gtg cgc gag 2229
Ala Gln Cys Ala Pro Pro Pro Ala Val Cys Ala Glu Leu Val Arg Glu
50 55 60

ccg ggc tgc ggc tgc tgc ctg acg tgc gca ctg agc gag ggc cag ccg 2277
Pro Gly Cys Gly Cys Cys Leu Thr Cys Ala Leu Ser Glu Gly Gln Pro
65 70 75 80

tgc ggc atc tac acc gag cgc tgt ggc tcc ggc ctt cgc tgc cag ccg 2325
Cys Gly Ile Tyr Thr Glu Arg Cys Gly Ser Gly Leu Arg Cys Gln Pro
85 90 95

tcg ccc gac gag gcg cga ccg ctg cag gcg ctg ctg gac ggc cgc ggg 2373
Ser Pro Asp Glu Ala Arg Pro Leu Gln Ala Leu Leu Asp Gly Arg Gly
100 105 110

ctc tgc gtc aac gct agt gcc gtc agc cgc ctg cgc gcc tac ctg ctg 2421
Leu Cys Val Asn Ala Ser Ala Val Ser Arg Leu Arg Ala Tyr Leu Leu
115 120 125

cca gcg ccg cca gct cca g gtgagccgcc cgccaggtgc gctgcgtgca 2470
Pro Ala Pro Pro Ala Pro
130

gcaccgccac tggcgccgaa gggcctgggg gttgctgggt gccgctgcgg gagactccgc 2530

ttttcttctc actggagata atatgtgggg aaactgaagg cgctccggga aaggtgaagg 2590

cggtcgccga gggaccctcc ccagccggcc ctctacttgc tcgattctct aagtgcagag 2650

tacttgtaaa ttgcaaagcg ctttcagtga aaatgggtaa aggtttccgg agctgagggg 2710

agcggtaccg atgtttagct gttggaaaga tcctggacac aggagattct cctcgccccg 2770

cacgggtgca cacggactgc aatcccaggg atgcttgggg atgggggggat ataggcggat 2830

ttggaccaag gaaggtgggt aggcacgttg taggaaatag tacctctctt ttaaaatact 2890

gactttgcac agccttttgg tttgcaaagc aatgtctagt cccggtatgt ccaaaaacaa 2950

gtaaagtgga ttcgggtttt gatatcttct gcggttggaa aacctgaagc tgaaaaagaa 3010

gtaacttctt aaggttaccc agcggccaca acagagtgta ggtttgaact ccgcgtgcca 3070

ctttcagtac cataccattc ttacaactcg ggccacccct gcacctgcgc cgacctcaaa 3130

caaacttcca ggtgcgtggt gggtgcgggc aatgtggact aagtcaattt caatgacacg 3190

gcaagggaat tggaatcagt cctaggctgt ctcccttctt aatctgaaat gggggggggg 3250

aatgagatgt tgttaagggg agccccagaa gaggaaaaat gcaaacattt ggcagagtta 3310

ccctcttgct tagccactat cagtatcagg cagacagcga ctctggtaag ggcatcacat 3370

tgttcccttа aaaaaaggag cgggggttgt ttaaatggat ttggcagctg ttctttcaag 3430

cattcttagc cagcctcacc tagttatatg agaaataaag ttcctgcctt gcacagctga 3490

aggctgggag aattctcccc atcctaattc ccccaactcc ccaacgatca cgttggacag 3550

atgtcactgg gcaggccccc atctagggct agcaggatga acagtccctt tataatttat 3610

gtagctgtag agttccacgc ccgggtgaag ttattttctg gctcggcaag gctggctctg 3670

ttcacccctg agaaatgctg gattcatgga aaggcaagat gcctgaaaca tacactggct 3730

ctggtcagct gttaaagctg ctggaggcat ttgtctctcg gggcaaagtt atgtcatttg 3790

ccaagtgtcg tacattattg tgcattttgg ggtattcaaa aagtgatctt agaaatactg 3850

atacacatcg tcattcttgg gctttagcaa tcatcatgat taccacctta gtagcactgt 3910

agtataggtt gatgtgagtt ataagattat aaaaagatct aagtgacttc tagaatctat 3970

ttgacaaaaa aaggtaaatt ttcgacagtc aaaagtcaca attatctgtt gcttaaatag 4030

aactgttttg tcttcatgcc ctagtctgca gcccaggcat taagaagaaa ccaaggaaat 4090

```
ttaagaaatt actcaaggtt cttagaaaag aagtataaat acgtttattt acatgttctt   4150

agagtattta cattcttagt atctctttta tctcagtatt tccttgaaaa agaaagcaag   4210

ctaagattaa aagaaattga aaccaaatcc tcgcaggtag ggacctcctc tgtgaggctc   4270

tgtgctggac cctgggaatg tgtgcttccc aaggtatgaa acccccttggg gaactttaca   4330

gcaggacctc agtgagctgt ttggcaggtg aggaaactaa gacccagaga ggagagggac   4390

tttcctaagg ccctggtgag tgacctgcca gtagccactt ccaggggaga gcagagcatc   4450

tgcagccaaa tcattgcagc cccaggtagc tttctagata gactgtggac cagatgggcc   4510

acctgagctc cctgctaggg ttacacatta tagccctgtt tgtgtagtag agaaatttca   4570

tgactctcaa ttgtggactt aagccgatgc ctccagacct tggcatggtc cacaggccct   4630

gggagcatgg gctctgaatg tagcctttga tccccatagc ggtcttacag cccctccaag   4690

ttcattctga agaaggaatg gagtgagaat cctggctgca gatccagtct tgaatttagt   4750

catatactta aaattccaat tcaactgtta acattccagc atccatttta agcatcagac   4810

tttcttcatt tagcactttt tattataaaa gggagatctg ctggaggggg atttctccta   4870

ccccaccccc acccagggaa ggaaaagctc tttggcactt agaagtctga gccgtgagtg   4930
attgtctgca tccatgtgct gctgtgttca cccggggtga aaaggactca   4990

cttaggcagg caccagcaag atgcacaggg tctgtgtaga ccttgagttt tagagatgta   5050

acggggacct agaaaacaag ccaccaacat gcttgcatga ttctgagccc ctgaggcaaa   5110

acgctttgca ggtaataatt cagttttccc atctgagctg gacaccaagc tcttataagc   5170

gtgtttacct ggtagcattg aggacggtac tggtcaacct tggaattccc ataagggctt   5230

gttacaactc agactcgtgc cgccactcca gcgtttccgg agtggagaat gtgcatttct   5290

tccaagtccc cgggctgccg ctgctcccgc gggtgggagg accacacttg gagttgactg   5350

caaaatttct gagccggcgc tgcagcagcc tcccgtggct caggtctgcc ccctgccggt   5410

ggaagatgaa gcatactgcc ttcacctact gaggggcact gaagcgtttg tctgccttct   5470

ttagttgcag ctacttagga agagcacctg tcagattgac tttcaaacag ataacttctt   5530

gaggtagagc aaccaccatg tagtgagtag tatgatggaa taatacttca tcgaggtatt   5590

taaaaaaaaa acctcacttg gattgccaac taatattgtc atttacatgt gacctggttg   5650

caacgttaag atttttacaa gactgtgata gatattgatg actctcatgt gtttgtctct   5710

cttgggcgtt ttaag ga  aat gct agt gag tcg gag gaa gac cgc agc gcc   5760
                Gly Asn Ala Ser Glu Ser Glu Glu Asp Arg Ser Ala
                                    140                 145

ggc agt gtg gag agc ccg tcc gtc tcc agc acg cac cgg gtg tct gat   5808
Gly Ser Val Glu Ser Pro Ser Val Ser Ser Thr His Arg Val Ser Asp
            150                 155                 160
```

```
ccc aag ttc cac ccc ctc cat tca aag ata atc atc atc aag aaa ggg      5856
Pro Lys Phe His Pro Leu His Ser Lys Ile Ile Ile Ile Lys Lys Gly
        165                 170                 175

cat gct aaa gac agc cag cgc tac aaa gtt gac tac gag tct cag agc      5904
His Ala Lys Asp Ser Gln Arg Tyr Lys Val Asp Tyr Glu Ser Gln Ser
    180                 185                 190

aca gat acc cag aac ttc tcc tcc gag tcc aag cgg gag aca gaa tat      5952
Thr Asp Thr Gln Asn Phe Ser Ser Glu Ser Lys Arg Glu Thr Glu Tyr
195                 200                 205                 210

gtgagagctt ttcctcttgt taaaggagga gggcaagacc tgccaagcct gggtactcag    6012

agcctcttga gggcaattct tactcaacaa accccagcgc ctggctgatg ggtgggcaac    6072

ccctagcccc tctgtgccct acctctctcc tctccttaca taaagaatat tgacccttt     6132

ggagaatctt atgaggatca agctgaaata acactcttaa aagcatatgg gatgtcataa    6192

agacctctgc agataatgaa aatattctca taaagatagt tttatttact tcatcctcta    6252

tgcttgttga cctgctattg gttccatgcc agcttctgtg ccttactctg ggaagagcaa    6312

aaaggagaca gggagtgatg gttagcttat tcgggggact ttcgtgctac atcagacata    6372

aggtatctga ggagcaaatt acaggtccca cttttggtag ttgtgcagca tcgtaagatt    6432

tttaaagcac acattctaga gtaaaaactg tgactctgtt gctctggtcc ttcctgatcc    6492

ccag ggt ccc tgc cgt aga gaa atg gaa gac aca ctg aat cac ctg aag      6541
     Gly Pro Cys Arg Arg Glu Met Glu Asp Thr Leu Asn His Leu Lys
                    215                 220                 225

ttc ctc aat gtg ctg agt ccc agg ggt gta cac att ccc aac tgt gac      6589
Phe Leu Asn Val Leu Ser Pro Arg Gly Val His Ile Pro Asn Cys Asp
                230                 235                 240

aag aag gga ttt tat aag aaa aag cag gtgagtgagg tcctcagtgt            6636
Lys Lys Gly Phe Tyr Lys Lys Lys Gln
            245                 250

gttttcttcc tcttctgttg acacagagga gaaacccatg tcaccagcgc ccaggctctt    6696

gtggccatag ctctaactct gagcctgtgc agcaccagtg cccaggactt ggtgccagtc    6756

tcaggaggtc agaccaaggg ctgctttgac ttgttgctct gagtgctgct atattggcca    6816

taatcctcaa ccctagtgcc tttccaccac ccgcttccca ctcctgtcct ttcaatggtt    6876

cacccacagg cggacaagat gctgcccagt ggcacccttt ataaactgca agtggacatg    6936

ttaacacatt tgttaatgct gcgtcaggga gtgacatttc aaacaactat tatagtcagt    6996

ttccaagaag tgtgacatga ggtcatacca caaaaaagct taccctgaaa tcccacaatc    7056

gtccccttc ctactgatgc cttcccgata gtgagcaggt tgcaatatta agattttgaa     7116

aaggctgttg ctagatgttg gtgactcgtg tgtctctgtc tcccttgggc ttttcaagga    7176

aatgctagtg agtgggggga tgactgcagc atggccagct tggagagccc agccatcccc    7236
```

agcacatacc aggtgtctgt cttggcgtgg aggggatgga acttgaaatc agacactcgg 7296

tccatgctgg ggatggccag tctctccaaa ctggcatgtg gtcttcctcc gagtcactgg 7356

catttcccta gaaagtccaa gtgagaagaa ggcatgagag tcatcaacat caaacaacag 7416

tcttttcaaa atctttatat tgcaacatag tcccattcct ggaaaaggaa tggagtgaga 7476

atcctggcta cacatcagcc ccaaatgtag tcattgccta aaatcccaat taacctgaaa 7536

atgatcaaac aaatttaaga tatagtaata ttaagctgta ataaatatgc ttctataggc 7596

tttgtgttat gtgatggcac tatttcaatt ggctttctaa ttggacaatt gatactatgc 7656

tatctacaga attggccttt ggagacctaa gtgagccaca gtggcctcag ggtgaccata 7716

tactaggatt catagcagtg gccacagtca gaagcctaag ctttcctcca ttgccattgc 7776

tcgtttatac cacgtttctg tcaaagtcat attcattcaa caaagtcata ctgagaaggt 7836

gtcatgtgag gctggatgtg ggctccaaag tcatagctgt gacattcgca ggcagcggga 7896

tgttctcagt tccacatttg gcagagaagt cagtcaagag gttctacaag ggctggtgtc 7956

caccttatac tcctagaaac acaaaactgc ccccaccccc gctttcttgg agcaggaagt 8016

tacacccaca cgcatgcaca ggcgcacact cagcgggcct aggcagcgtg gctcttgtgt 8076

tgccttagct gaaatttctg ttgtgctttc tcagcatagc agagtcacgc tggcaaacca 8136

tcatgcgccc tggccaccga cctgacacca gacccaggag cattcacttc tctgtcttct 8196

gtttctctcc cacag tgt cgc cct tcc aaa ggc agg aag cgg ggc ttc tgc 8247
Cys Arg Pro Ser Lys Gly Arg Lys Arg Gly Phe Cys
255 260

tgg tgt gtg gat aag tat ggg cag cct ctc cca ggc tac acc acc aag 8295
Trp Cys Val Asp Lys Tyr Gly Gln Pro Leu Pro Gly Tyr Thr Thr Lys
265 270 275

ggg aag gag gac gtg cac tgc tac agc atg cag agc aag tagacgcctg 8344
Gly Lys Glu Asp Val His Cys Tyr Ser Met Gln Ser Lys
280 285 290

ccgcaagggt gagtactcag gaggggcagc ctgggctcca gggcctcact gtccttggac 8404

cagcctcagg ggctgggcgt ggccactggc cttccccagg cttacagacc caggagctgc 8464

agctcagggc cagaaagagc aaagcaaata ggacagagcc ctcagaaggg tgcagggaga 8524

gggagacccc atcaacccaa ccaaacaagt gtggggaagg aggccggcca gtgcacctca 8584

gggacactct gctttatctc agatacctca cagcacctaa gctatcattc atccacacac 8644

aaagtgaaga ttttcaaagt taggctttac ccgtgagtct ggaggtcatt tatcttcaca 8704

gagaacgttt atcgcagact gctaagatac atgttctaat taagatgtga tgtgagaacg 8764

ctgaatgctc gttggagact cagttgaagt gcagcttttt ttctgtcaaa tatataatga 8824

atattctgtt agtctgtggc taatataatt ttaataaagt taatttaaat ctgatagaaa 8884

aatgaaattt taaacgataa ttttagagaa tgctattata tccagtcttc ttttttcttt 8944

taataaatga gggaactatt gggggaaagg aataaataca ttttctttca ttttattaag 9004

acaaatttag taagcagaag aaatttgcat gtttagttat aagggtttct tttttcctta 9064

caagttggaa aaaataattc taatttaagg gtaactcttt gacaatgaac actgtgagca 9124

gcatctggta ctcgttgctt tgtttgaaaa catgagttga gaccccagcc gcacttgcag 9184

cctagtgcca ttagcctgca ggctgtgctg gatatctcag ggcaagagtc gagccctttt 9244

gattttgggg ggattatttc aatatatttg ctttttcttt ttgttttagt taatgtggag 9304

ctcaaatatg ccttattttg cacaaaagac tgccaaggac atgaccagca gctggctaca 9364

gcctcgattt atatttctgt ttgtggtgaa ctgatttttt ttaaaccaaa gtttagaaag 9424

aggtttttga aatgcctatg gtttctttga atggtaaact tgagcatctt ttcactttcc 9484

agtagtcagc aaagagcagt ttgaattttc ttgtcgcttc ctatcaaaat attcagagac 9544

tcgagcacag cacccagact tcatgcgccc gtggaatgct caccacatgt tggtcgaagc 9604

ggccgaccac tgactttgtg acttaggcgg ctgtgttgcc tatgtagaga acacgcttca 9664

cccccactcc ccgtacagtg cgcacaggct ttatcgagaa taggaaaacc tttaaacccc 9724

ggtcatccgg acatcccaac gcatgctcct ggagctcaca gccttctgtg gtgtcatttc 9784

tgaaacaagg gcgtggatcc ctcaaccaag aagaatgttt atgtcttcaa gtgacctgta 9844

ctgcttgggg actattggag aaaataaggt ggagtcctac ttgtttaaaa aatatgtatc 9904

taagaatgtt ctagggcact ctgggaacct ataaaggcag gtatttcggg ccctcctctt 9964

caggaatctt cctgaagaca tggcccagtc gaaggcccag gatggctttt gctgcggccc 10024

cgtggggtag gagggacaga gagacgggag agtcagcctc cacattcaga ggcatcacaa 10084

gtaatggcac aattcttcgg atgactgcag aaaatagtgt tttgtagttc aacaactcaa 10144

gacgaagctt atttctgagg ataagctctt taaaggcaaa gctttatttt catctctcat 10204

cttttgtcct ccttagcaca atgtaaaaaa gaatagtaat atcagaacag gaaggaggaa 10264

tggcttgctg gggagcccat ccaggacact gggagcacat agagattcac ccatgtttgt 10324

tgaacttaga gtcattctca tgcttttctt tataattcac acatatatgc agagaagata 10384

tgttcttgtt aacattgtat acaacatagc cccaaatata gtaagatcta tactagataa 10444

tcctagatga aatgttagag atgctatatg atacaactgt ggccatgact gaggaaagga 10504

gctcacgccc agagactggg ctgctctccc ggaggccaaa cccaagaagg tctggcaaag 10564

tcaggctcag ggagactctg ccctgctgca gacctcggtg tggacacacg ctgcatagag 10624

ctctccttga aaacagaggg gtctcaagac attctgccta cctattagct tttctttatt 10684

tttttaactt tttgggggga aaagtatttt tgagaagttt gtcttgcaat gtatttataa 10744

atagtaaata aagtttttac cattaaaaaa atatctttcc ctttgttatt gaccatctct 10804

gggctttgta tcactaatta ttttatttta ttatataata attattttat taaaatgttc 10864

cctgctttcc cttttagcaa 10884

<210> 12
<211> 291
<212> PRT
<213> Homo sapiens

<400> 12

Met Gln Arg Ala Arg Pro Thr Leu Trp Ala Ala Ala Leu Thr Leu Leu
1 5 10 15

Val Leu Leu Arg Gly Pro Pro Val Ala Arg Ala Gly Ala Ser Ser Gly
20 25 30

Gly Leu Gly Pro Val Val Arg Cys Glu Pro Cys Asp Ala Arg Ala Leu
35 40 45

Ala Gln Cys Ala Pro Pro Pro Ala Val Cys Ala Glu Leu Val Arg Glu
50 55 60

Pro Gly Cys Gly Cys Cys Leu Thr Cys Ala Leu Ser Glu Gly Gln Pro
65 70 75 80

Cys Gly Ile Tyr Thr Glu Arg Cys Gly Ser Gly Leu Arg Cys Gln Pro
85 90 95

Ser Pro Asp Glu Ala Arg Pro Leu Gln Ala Leu Leu Asp Gly Arg Gly
100 105 110

Leu Cys Val Asn Ala Ser Ala Val Ser Arg Leu Arg Ala Tyr Leu Leu
115 120 125

Pro Ala Pro Pro Ala Pro Gly Asn Ala Ser Glu Ser Glu Glu Asp Arg
130 135 140

Ser Ala Gly Ser Val Glu Ser Pro Ser Val Ser Ser Thr His Arg Val
145 150 155 160

Ser Asp Pro Lys Phe His Pro Leu His Ser Lys Ile Ile Ile Ile Lys
165 170 175

Lys Gly His Ala Lys Asp Ser Gln Arg Tyr Lys Val Asp Tyr Glu Ser
180 185 190

```
Gln Ser Thr Asp Thr Gln Asn Phe Ser Ser Glu Ser Lys Arg Glu Thr
        195                 200                 205

Glu Tyr Gly Pro Cys Arg Arg Glu Met Glu Asp Thr Leu Asn His Leu
    210                 215                 220

Lys Phe Leu Asn Val Leu Ser Pro Arg Gly Val His Ile Pro Asn Cys
225                 230                 235                 240

Asp Lys Lys Gly Phe Tyr Lys Lys Lys Gln Cys Arg Pro Ser Lys Gly
                245                 250                 255

Arg Lys Arg Gly Phe Cys Trp Cys Val Asp Lys Tyr Gly Gln Pro Leu
            260                 265                 270

Pro Gly Tyr Thr Thr Lys Gly Lys Glu Asp Val His Cys Tyr Ser Met
        275                 280                 285

Gln Ser Lys
    290
```

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/00558

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. [7]: C12Q 1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

SEE ELECTRONIC DATA BASES

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

SEE ELECTRONIC DATA BASES

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

Medline, CA:IGFBP, insulin like growth factor binding protein, liver, hepato, HCC, cancer, tumour, tumor, neoplasm, carcinoma, malignant

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99/46597 A1 (DIAGNOSTIC SYSTEMS LABORATORIES INC) 16 September 1999 Whole Document	1 - 18
P, X	WO 01/53837 A1 (DIAGNOSTIC SYSTEMS LABORATORIES INC) 26 July 2001 Whole Document	1 - 18
X	WO 00/69454 A1 (BOARD OF REGENTS, THE UNIVERISTY OF TEXAS SYSTEM) 23 November 2000 Whole Document	27 - 29

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
16 May 2002	12 JUN 2002
Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	**CRAIG ALLATT** Telephone No : (02) 6283 2414

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/00558

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JG Reeve et al *Int J Cancer* (1992) 51(5) pp 818 - 21 "Insulin-like growth-factor-binding protein gene expression and protein production by human tumour cell lines" Whole Document	1 - 18
X	SA Weinzimer et al *J Clin Endocrinol Metab* (2001) 86(4) pp 1806 - 13. "Transferrin is an insulin-like growth factor-binding protein-3 binding protein" Whole Document	24
X	T Yamashita et al *Biochem Biophys Res Commun* (2001) 282(2) pp 647 - 54. "Serial analysis of gene expression in chronic hepatitis C and hepatocellular carcinoma" Whole Document	1 - 18
X	N Kondoh et al *Cancer Res* (1999) 59(19) pp 4990 - 6 "Identification and characterization of genes associated with human hepatocellular carcinogenesis" Whole Document	1 - 18
X	JG Scharf et al *Lab Invest* (2000) 80(9) pp 1399 - 411 "Analysis of the IGF axis in preneoplastic hepatic foci and hepatocellular neoplasms developing after low-number pancreatic islet transplantation into the livers of streptozotocin diabetic rats" Whole Document	1 - 18
X	Y Gong et al *Mol Cell Biochem* (2000) 207(1-2) pp 101 - 4 "The expression of insulin-like growth factor binding proteins in human hepatocellular carcinoma Whole Document	1 - 18
X	M van Kleffens et al *Endocrinology* (1999) 140(12) pp 5944 - 52 "Generation of antisera to mouse insulin-like growth factor binding proteins (IGFBP)-1 to -6: comparison of IGFBP protein and messenger ribonucleic acid localization in the mouse embryo" Whole Document	24
X	Y Shirota et al *Hepatology* (2001) 33(4) pp 832 - 40 "Identification of differentially expressed genes in hepatocellular carcinoma with cDNA microarrays" Hepatology" Whole Document	1 - 18
X	SG Gray et al" *Br J Cancer* (2000) 82(9) pp 1561 - 7 "Altered expression of members of the IGF-axis in hepatoblastomas" Abstract, pp 1565 - 7	1 - 18
X	H von Horn et al *Cancer Lett* (2001) 162(2) pp 253 - 60 "Expression levels of insulin-like growth factor binding proteins and insulin receptor isoforms in hepatoblastomas" Whole Document	1 - 18
X	L Xu et al. *Cancer Res* (2001) Apr 1 61(7) pp 3176 - 81 "Expression profiling suggested a regulatory role of liver-enriched transcription factors in human hepatocellular carcinoma" Figure 5	1 - 18

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU02/00558

Box I Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos :

because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos : **19 - 23, 25, 26, 30 - 37**

because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

The claims are not limited to the technical features that define the invention.

3. ☐ Claims Nos :

because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a)

Box II Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Claims 1 - 18 are directed toward a method for detecting the presence of liver cancer by detecting the aberrant expression of a gene encoding an insulin like growth factor binding protein (IGFBP).

Claims 19 - 37 are directed toward methods of treating liver cancer by modulating expression of an insulin like growth factor binding protein.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/AU02/00558

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member			
WO	99/45697	EP	1066716	SE	9800686
WO	2001/53873	FR	2803918		
WO	2000/69454	NONE			

END OF ANNEX